SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Mihaela Poljak

TEORIJA POVJERENJA

Diplomski rad

Voditelj rada:
prof. dr. sc. Miljenko Huzak

Zagreb, srpanj 2014.

Ovaj diplomski rad obranjen je dana ____________________ pred ispitnim povjerenstvom u sastavu:

1. ____________________________, predsjednik
2. ____________________________, član
3. ____________________________, član

Povjerenstvo je rad ocijenilo ocjenom ________________.

Potpisi članova povjerenstva:

1. ____________________________
2. ____________________________
3. ____________________________

Zahvaljujem svom mentoru, prof. dr. sc. Miljenku Huzaku na strpljenju, savjetima i pomoći pri izradi diplomskog rada.
Hvala mojoj obitelji na velikoj podršci i razumijevanju tokom svih ovih godina studiranja.

Sadržaj

Uvod

Kao i igrama na sreću, i osiguranjem prenosimo rizik i dobitak (ili gubitak). Ali tu je sličnostima kraj. Osiguranje je jedan oblik upravljanja rizikom, prvenstveno usmjeren na smanjenje financijskih gubitaka. Osnovna ideja koja leži iza osiguranja je da osiguranici, bili to pojedinci, grupe ili kompanije uplaćuju premije osiguranja i na taj način prenose svoj rizik na osiguravajuće društvo.

Glavni cilj svakog osiguravajućeg društva je biti što konkurentniji na tržištu. Da bi to postigli od velike važnosti je odrediti što točniju individualnu premiju za rizik. Ideja je svrstati pojedince koji su podjednako izloženi riziku u istu grupu. Tu grupu tada zovemo kolektiv. Na primjer, pojedinci znaju da je rizik požara u njihovoj kući mali, ali ga se boje zato što bi posljedice bile veoma nepovoljne. Stoga su spremni platiti premiju osiguranja protiv požara. Osiguravajuća društva primaju uplatu premija zašto što znaju da će se dogoditi mali broj požara unutar velikog kolektiva. Stoga je i osiguravajuća premija za požare obično niska.

U ovom diplomskom radu bavimo se problemom određivanja što točnije individualne premije za rizik. Međutim, nije jednostavno odrediti homogene grupe rizika. Na veličinu premije utječe velik broj faktora, stoga se mora voditi posebna pažnja da grupe ne bi bile premale jer je tada premalo statističkih podataka na osnovu kojih bi trebali procijeniti premiju rizika. S druge strane problem kod prevelikih grupa bio bi nehomogenost grupe što bi vodilo neadekvadnoj procjeni pojedinih rizika. Zbog toga do izražaja dolazi teorija povjerenja. Ona je matematički alat u rješavanju problema heterogenog kolektiva. Odgovara na pitanje kako kombinirati iznose pojedinačnih šteta s prosječnom štetom sličnih rizika, tako da procjena premije bude što je moguće bolja. Ova teorija je jedan od osnovnih alata u modernom aktuarstvu, a matematički pripada području Bayesove statistike.

Dakle, u prvom poglavlju se upoznajemo s temom kroz matematičku formulaciju problema procjene rizika u osiguranju, tj. kroz definicije točne individualne, kolektivne i Bayesove premije te njihovim međusobnim odnosima.

U drugom poglavlju pokazujemo zašto je Bayesova premija najbolja procjena premije. Promatram ju u tri specijalna modela: model Poisson-gama distribucija, model binomna-beta distribucija i model normalna-normalna distribucija. Unatoč vrlo elegantno izvedenom teorijskom rješenju računanja premije provedenom u ovom poglavlju dolazimo do problema primjene te teorije u praksi. Općenito premiju ne možemo egzaktno izračunati jer može biti vrlo komplicirana. Iz tog razloga u trećem poglavlju umjesto Bayesovog procjenitelja, procjenitelj za individualnu premiju tražimo među linearnim funkcijama našeg uzorka. Takve procjenitelje ćemo zvati procjenitelji povjerenja i tražiti ćemo ih u raznim modelima, što nas dovodi do četvrtog poglavlja a to je Bülhmann-Straubov model.

Poglavlje 1

Matematička formulacija problema procjene rizika u osiguranju

1.1 Individualni rizik i točna individualna premija

Neka je sa X_j, označen ukupan iznos šteta za promatranog pojedinca u godini j ili u nekom drugom dobro definiranom periodu j, $j = 1, ..., n$. Veličine $X_j, j = 1, ..., n$, matematički interpretiramo kao slučajne varijable.

Na temelju podataka o štetama pojedinca u prethodnim periodima, $\mathbf{X} = (X_1, ..., X_n)'$, želimo odrediti njegovu premiju rizika ukupne štete za sljedeći period, naime, želimo procijeniti X_{n+1}, tj. njegovu očekivanu štetu u sljedećem periodu. U tu svrhu moramo imati određene pretpostavke za funkciju distribucije $F(x)$ slučajne varijable X_j.

Standardne pretpostavke funkcije distribucije su:

A1: stacionarnost: Sve slučajne varijable X_j su jednako distribuirane s (uvjetnom) funkcijom distribucije $F(x)$.

A2: (uvjetna) nezavisnost: Slučajne varijable $X_j, j = 1, 2, ...,$ (uvjetno) nezavisne (za danu funkciju distribucije $F(x)$).

Napomena

- Kasnije u radu ćemo precizno objasniti što mislimo pod uvjetnom nezavisnošću slučajnih varijabli za danu uvjetnu funkciju distribucije $F(x)$.

Pretpostavaka A1 potrebna nam je da bismo ustanovili vezu između prošlosti i budućnosti. Odnosno, potrebna nam je da bismo izračunali premiju osiguranja na temelju podataka u prethodnim periodima.

U osiguravajućoj praksi često se nalazimo u situaciji da je funkcija distribucije F nepoznata i da varira među rizicima. U matematičkoj statistici za F koristimo oznaku F_ϑ, gdje je ϑ nepoznat parametar i varira među rizicima. ϑ zovemo *profil rizika* (eng. *risk profile*), a općenito uzimamo da je ϑ element nekog apstraktnog skupa Θ.

Princip računanja premije je funkcija koja slučajnoj varijabli X pridružuje realan broj $H(X)$:

$$X \mapsto H(X).$$

Neki dobro poznati principi za računanje premije su:

- PRINCIP OČEKIVANJA: $X \mapsto (1+\alpha)E[X] \quad \alpha > 0$;
- PRINCIP STANDARDNE DEVIJACIJE: $X \mapsto E[X] + \beta\sigma(X) \quad \beta > 0$;
- PRINCIP VARIJANCE: $X \mapsto E[X] + \gamma\sigma^2(X) \quad \gamma > 0$;
- EKSPONENCIJALNI PRINCIP: $X \mapsto \frac{1}{\delta}\ln E[e^{\delta}X] \quad \delta > 0$.

Za svaki od tih principa, premiju $H(X)$ možemo raščlaniti na čistu premiju rizika (eng. *pure risk premium*) $E[X]$ i pozitivni ostatak određen parametrima α, β, γ i δ. Primjećujemo da čistu premija rizika $E[X]$ možemo dobiti principom očekivanja gdje je $\alpha = 0$. Kako bi svaka osiguravajuća kompanija uspjela svladati nestalnost i u nepovoljnim godinama za nju ispuniti sve obaveze prema svojim klijentima, mora imati pridržaj. Taj pridržaj predstavlja pozitivni ostatak. U njega ulazi i naknada investitorima za izlaganje svog kapitala riziku, koju promatramo kao trošak rizičnog kapitala.

Primjenom principa H, za rezultat dobijemo točnu individualnu premiju za rizik s profilom rizika ϑ, što ćemo označiti sa $H(F_\vartheta)$. Ograničit ćemo se na promatranje čiste premije rizika što nas dovodi do sljedeće definicije čiste premije rizika.

Definicija 1.1.1. *Točna individualna premija (eng. correct individual premium) za rizik s profilom rizika ϑ je*

$$P^{ind}(\vartheta) = E[X|\vartheta] = \mu(\vartheta) := \int x \,\mathrm{d}F_\vartheta(x). \tag{1.1}$$

Točnu individualnu premiju još nazivamo i fer premija rizika. Ubuduće, radi jednostavnosti u zapisu, umjesto $E[\cdot|F_\vartheta]$ ili $E[\cdot|\vartheta]$ koristit ćemo oznaku $\underline{E_\vartheta}[\cdot]$. U praksi osiguranja, i ϑ i $\mu(\vartheta)$ su nepoznate veličine, stoga ćemo tražiti procjenitelj $\widehat{\mu(\vartheta)}$ za $\mu(\vartheta)$.

1.2 Procjena rizika i premija kolektiva

Osiguravajuće kompanije osiguravaju različite vrste rizika. Rizici su grupirani u razrede "sličnih rizika", na bazi tzv. "objektivnih" karakteristika rizika. Na primjer, neke karakteristike koje promatramo kod osiguranja automobila su: kapacitet cilindra, vrsta, snaga, težina motora te karakteristike vezane za samog vozača, kao što su starost, spol, regija i slično. U osiguranju od požara, neke bitne karakteristike bi bile tip konstrukcije osigurane građevine, vrsta posla koji se obavlja u tom objektu ili mogućnost gašenja požara u objektu. U praksi je vrlo važno kako se te grupe formiraju, ali u ovom radu nećemo se baviti tom konstrukcijom.

Bitna značajka teorije povjerenja jest da svaki pojedini rizik ne promatramo individualno, nego kao dio grupe sličnih rizika kojoj pripada, što zovemo kolektiv.

Taj problem formuliramo na sljedeći način:

Svaki rizik i iz kolektiva opisan je svojim individualnim profilom rizika ϑ_i. Parametri ϑ_i su elementi skupa Θ, gdje je Θ skup svih mogućih vrijednosti profila rizika za taj kolektiv.

Jasno je kako su rizici u kolektivu različiti, stoga se njihove ϑ-vrijednosti međusobno razlikuju. Međutim, rekli smo da je kolektiv grupa sličnih rizika, a njihova se sličnost očituje u činjenici da se ϑ-vrijednost za svaki rizik iz tog kolektiva bira iz istog skupa Θ.

Specijalno, za homogeni kolektiv, skup Θ sadrži samo jedan element. Tada bi svaki član kolektiva imao isti profil rizika, odnosno svi rizici u kolektivu bi bili jednaki. Međutim, u praksi to nije slučaj.

Iako su ϑ-vrijednosti određenih rizika u kolektivu nepoznate osiguravateljima, ipak na temelju *a priori* spoznaja i statističkih podataka mogu se izvući neke informacije o strukturi kolektiva. To je npr. podatak da većina vozača predstavlja "dobre" rizike jer vrlo rijetko čine štete dok samo mali postotak vozača čini učestale štete. Formalno, takve informacije mogu se sažeti vjerojatnosnom distribucijom $U(\vartheta)$ na skupu Θ.

Definicija 1.2.1. *Vjerojatnosnu funkciju distribucije $U(\vartheta)$ zovemo strukturna funkcija kolektiva.*

Interpretacija $U(\vartheta)$:

- *empirijska Bayesova interpretacija*: ako ϑ u kolektivu promatramo kao slučajni uzorak iz fiksnog skupa Θ, onda funkcija $U(\vartheta)$ opisuje frekvenciju pojavljivanja ϑ na skupu Θ.

- *čista Bayesova interpretacija*: funkciju distribucije $U(\vartheta)$ promatramo kao opis osobnog mišljenja, a priori spoznaja i iskustva aktuara.

Definicija 1.2.2. *Kolektivna premija (eng. collective premium) definirana je sa:*

$$P^{coll} = \int_{\Theta} \mu(\vartheta) \mathrm{d}U(\vartheta) =: \mu_0. \tag{1.2}$$

Komentar promatranih premija

- Točna individualna premija $P^{ind}(\vartheta) = \mu(\vartheta) = E_{\vartheta}[X_{n+1}]$ odgovara očekivanom iznosu štete individualnog rizika (za dani profil rizika ϑ) za period $n+1$. Budući da je za osiguravatelja parametar ϑ, pa onda i $\mu(\vartheta)$ nepoznat, moramo procijeniti premiju. Međutim, u najboljem slučaju, osiguravatelj ima na raspolaganju informaciju o štetama za pojedinog osiguranika za nekoliko prethodnih perioda. Ali, ta informacija je često dosta ograničena i nije najpouzdaniji pokazatelj očekivane vrijednosti u sljedećem periodu.

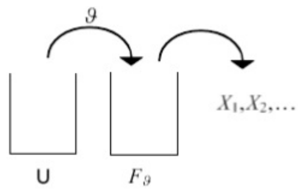
- Kolektivna premija $P^{coll} = \int_{\Theta} \mu(\vartheta) dU(\vartheta) =: \mu_0$ odgovara prosjeku očekivanih individualnih rizika za cijeli kolektiv. U većini slučajeva, ova vrijednost je nepoznata osiguravatelju, ali ipak za razliku od individualne premije, ovu premiju možemo procijeniti s vrlo velikom preciznošću na temelju podataka o prethodnim periodima.

Za osiguravajuću kompaniju od velike je važnosti mogućnost izračuna kolektivne premije. Budući da je grupa sličnih a ne istih rizika, ne možemo istu premiju odrediti za svakog člana kolektiva. Ali, prirodno je za očekivati da iznos premije za pojedinca treba ovisiti o njegovom ponašanju i o prosječnom ponašanju kolektiva jer je to ipak grupa sličnih rizika. Na primjer, ako se osiguranik u toj grupi osiguranja nalazi kratko vrijeme, onda je teško procijeniti njegovu očekivanu štetu za sljedeći period na temelju podataka o prethodim periodima.

U tržišnom nadmetanju osiguravajućih kompanija ključ uspjeha je ponuditi što bolju fer premiju rizika. Ako kompanija odredi jednaku premiju za sve rizike u heterogenom kolektivu, onda će dobri rizici platiti previše, a loši premalo. U tom slučaju, otvara se mogućnost da konkurentska kompanija ponudi niže premije za dobre rizike te će doći do slijevanja dobrih rizika u tu konkurentsku kompaniju. Prva kompanija bit će atraktivna samo za loše rizike pa zbog gubitaka svih dobrih rizika, može doći do pogubnih rezultata za tu osiguravajuću kuću.

1.3 Problem procjene premije na jeziku Bayesove statistike

Najelegantniji način za matematički opis problema procjene premije kolektiva je Bayesova statistika i to ćemo demonstrirati modelom pod nazivom *"model dvije urne"* (eng. *two-urn model*).



Slika 1.1: "Model dvije urne"

Prva urna predstavlja kolektiv sa funkcijom distribucije U. Iz te urne biramo individualni rizik, odnosno njegov profil rizika ϑ sa zakonom razdiobe $U(\vartheta)$. Parametar ϑ onda određuje sastav druge urne, odnosno funkciju distribucije F_ϑ. Zatim iz druge urne biramo vrijednosti slučajnih varijabli $X_1, X_2, ...$, koje su nezavisne i jednako distribuirane s funkcijom distribucije F_ϑ.

Matematički precizno opisujemo ovaj model na sljedeći način:

Svaki rizik je opisan svojim individualnim profilom rizika ϑ, a ϑ je realizacija slučajne varijable Θ i vrijedi:

- uvjetno na događaj $\Theta = \vartheta$, $X_1, X_2, ...$ su nezavisne jednako distribuirane slučajne varijable s funkcijom distribucije F_ϑ,
- Θ je slučajna varijabla s funkcijom distribucije U.

Par $(\Theta, (X_j))$, $j = 1, 2, ...$, predstavlja pojedinu policu u portfelju, gdje je parametar Θ nepoznat, a niz slučajnih varijabli (X_j) predstavlja niz šteta u danoj polici.

U daljnjem tekstu navodimo neke posljedice i zaključke vezane za naš model.

- Model *dvije urne* opisuje isti problem koji smo već opisivali. Rizici u kolektivu se međusobno razlikuju, svaki ima svoj profil rizika ϑ. Zajedničko im je to što u određenom kolektivu su njihovi profili ϑ_i realizacije iste slučajne varijable Θ. Vrijednosti ϑ_i su nezavisno odabrane iz iste urne s funkcijom distribucije U.

- U ovoj interpretaciji, individualnu premiju možemo promatrati kao slučajnu varijablu $\mu(\Theta)$. Ne znamo točnu vrijednost individualne premije, ali ipak nešto znamo o mogućim vrijednostima $\mu(\Theta)$, te o vjerojatnostima s kojima se te vrijednosti događaju, stoga je prirodno $\mu(\Theta)$ promatrati kao slučajnu varijablu. Dakle, individualnu premiju možemo definirati sa:

$$P^{ind} = \mu(\Theta) = E[X_{n+1}|\Theta],$$

  a to je uvjetno očekivanje štete X_{n+1} pa time i slučajna varijabla.

- Primijetimo da su svi rizici *a priori* jednaki. Znamo da postoje bolji i loši rizici u portfelju, ali ne možemo, *a priori*, znati kojem razredu pojedini rizik pripada. Određene zaključke o riziku možemo donositi tek *a posteriori*, dakle nakon promatranja individualnih rizika.

- Za razliku od individualne premije P^{ind} koja je slučajna varijabla, kolektivna premija

$$P^{coll} = \mu_0 = \int_{\Theta} \mu(\vartheta) dU(\vartheta) = E[X_{n+1}],$$

  je fiksni broj jer se radi o bezuvjetnom očekivanju štete X_{n+1}.

- Ranije smo rekli (poglavlje 1.1, pretpostavka A2) da su slučajne varijable $X_j, j = 1, 2, ...$ uvjetno nezavisne. To direktno slijedi iz našeg modela. $X_1, X_2, ...$ su uvjetno nezavisne za dano Θ, a Θ određuje F_ϑ. Bezuvjetno, $X_j, j = 1, 2, ...$ su pozitivno korelirane, tj. nisu i bezuvjetno međusobno nezavisne:

$$Cov(X_1, X_2) = E[Cov(X_1, X_2|\Theta)] + Cov(E[X_1|\Theta], E[X_2|\Theta]) =$$

$$= Cov(\mu(\Theta), \mu(\Theta)) = Var[\mu(\Theta)].$$

  Nezavisnost vrijedi u slučaju kada je $Var[\mu(\Theta)] = 0$.

Napomena

- Da bismo imali što manje oznaka, istu oznaku Θ koju smo ranije koristili kao oznaku za skup mogućih ϑ vrijednosti, sada ju koristimo za slučajnu varijablu čija je realizacija ϑ.

Naš cilj je za svaki pojedini rizik procijeniti točnu individualnu premiju $\mu(\Theta)$ što je preciznije moguće. Potencijalni procjenitelj je kolektivna premija μ_0. U tom slučaju premija bi bila procijenjena prosjekom očekivanih individualnih rizika za cijeli kolektiv. Taj procjenitelj je prikladan za novi rizik za kojeg nemamo nikakve podatke o štetama u prethodnim periodima. On uzima u obzir činjenicu da rizik pripada kolektivu čiji je profil rizika slučajno odabran iz "kolektivne" urne U, ali, ne uzima u obzir iskustvo individualnih šteta. Ako smo promatrali rizik tijekom razdoblja od n godina i ako vektor $\mathbf{X}$ označava iznose individualnih šteta u tom razdoblju, ta informacija treba doprinijeti procesu procjene. To nas vodi do sljedeće procjene premije:

Definicija 1.3.1. *Bayesova premija (="najbolja" premija) definirana je sa:*

$$P^{Bayes} = \widetilde{\mu(\Theta)} = E[\mu(\Theta)|\mathbf{X}]. \tag{1.3}$$

U sljedećem poglavlju pokazat ćemo kako je Bayesova premija P^{Bayes} najbolja premija koja se temelji na iskustvu individualnih šteta $\mathbf{X}$, te odgovoriti na pitanja zašto je "najbolja" premija.

Poglavlje 2

Bayesova premija

U ovom poglavlju proučavat ćemo najbolju premiju temeljenu na iskustvu, odnosno Bayse-ovu premiju koju smo definirali sa (1.3),

$$P^{Bayes} = \widetilde{\mu(\Theta)} = E[\mu(\Theta)|\mathbf{X}].$$

Pokazat ćemo u kojem smislu je P^{Bayes} "najbolja" i zašto ju zovemo Bayesova premija.

2.1 Bayesov rizik i procjenitelj

Prvo ćemo navesti neke osnovne elemente teorije statističkog zaključivanja koji će nam pomoći u rješavanju ovog problema.

Neka je zadan slučajni vektor $\mathbf{X} = (X_1, X_2, ..., X_n)'$. Funkcija distribucije

$$F_\vartheta(\mathbf{x}) := P_\vartheta[\mathbf{X} \leq \mathbf{x}]$$

je potpuno ili djelomično nepoznata (ekvivalentno: parametar ϑ je potpuno ili djelomično nepoznat). Zanima nas vrijednost funkcionala $g(\vartheta)$ parametra ϑ, stoga tražimo funkciju $T(\mathbf{X})$ koja ovisi samo o vektoru $\mathbf{X}$, a procjenjuje $g(\vartheta)$ "što je bolje moguće".

Koristit ćemo oznake:
$\vartheta \in \Theta$: Θ je skup svih parametara koji sadrže pravu vrijednost ϑ,
$T \in D$: D je skup svih funkcija kojem pripada i procjenitelj funkcije g.

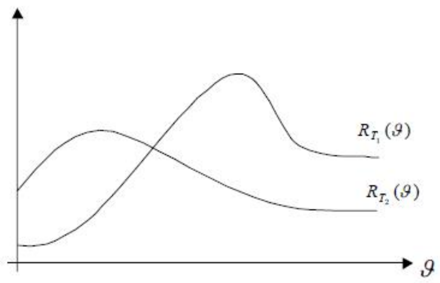
T je preslikavanje sa skupa $\mathbb{R}^n$ u skup svih mogućih vrijednosti funkcionala g, to jest, skup $\{g(\vartheta) : \vartheta \in \Theta\}$.

Da bismo precizno opisali značenje ideje *"što je bolje moguće"* uvodimo funkciju gubitka (eng. *loss function*) $L(\vartheta, T(\mathbf{x}))$ (kasnije ćemo definirati oblik za funkciju gubitka) te pomoću nje definiramo funkciju rizika (eng. *risk function*) za procjenitelj T:

$$R_T(\vartheta) := E_\vartheta[L(\vartheta, T)] = \int_{\mathbb{R}^n} L(\vartheta, T(\mathbf{x}))\, dF_\vartheta(\mathbf{x}).$$

Dozvoljeni su samo takvi T i L za koje desna strana postoji.

Cilj je pronaći takav procjenitelj $T \in D$, za koji je rizik $R_T(\vartheta)$ što je manji mogući. Međutim, općenito ne postoji procjenitelj T koji minimizira funkciju rizika $R_T(\vartheta)$ za sve ϑ. To ćemo ilustrirati primjerom na slici 2.1:



Slika 2.1: Funkcije rizika R_{T_1} i R_{T_2} za procjenitelje T_1 i T_2

Sada možemo definirati Bayesov rizik i Bayesov procjenitelj.

Definicija 2.1.1 (*Bayesov rizik*)**.** *Bayesov rizik za procjenitelj T u odnosu na a priori distribuciju $U(\vartheta)$ definiran je izrazom*

$$R(T) := \int_\Theta R_T(\vartheta)\, \mathrm{d}U(\vartheta).$$

Uz pretpostavku da definirani integral ima smisla, ovim kriterijem procjenitelje uvijek možemo urediti prema rastućem riziku, dakle skup svih procjenitelja $T \in D$ je uređen skup.

Napomena

Ako za realnu funkciju g definiranu na skupu S postoji $\widetilde{x} \in S$ tako da vrijedi $g(\widetilde{x}) \leq g(x)$, $\forall x \in S$, tada pišemo

$$\widetilde{x} := \arg\min_{x \in S}\ g(x).$$

Definicija 2.1.2. *Bayesov procjenitelj $\widetilde{T}$ definiramo kao*

$$\widetilde{T} := \arg\min_{T \in D_1} R(T),$$

gdje je D_1 skup svih dopustivih procjenitelja, a to su procjenitelji s integrabilnim funkcijama rizika.

Dakle, procjenitelj $\widetilde{T}$ je procjenitelj koji minimizira Bayesov rizik $R(\cdot)$. Bez daljnjeg razmatranja, uzet ćemo da je $\min_{T \in D_1} R(T)$ postignut.

Uvodimo sljedeće oznake:
P-zajednička distribucija vektora $(\Theta, \mathbf{X})$
F-marginalna distribucija od $\mathbf{X}$
$U_{\mathbf{X}}$-uvjetna distribucija od Θ, za dano $\mathbf{X} = \mathbf{x}$.

Da bismo konstruirali Bayesov procjenitelj, promatramo sljedeće jednadžbe koje dobijemo raspisom Bayesovog rizika $R(T)$ uz promjenu Fubinijevog teorema na nenegativnu podintegralnu funkciju $(\mathbf{x}, \vartheta) \mapsto L(\vartheta, T(\mathbf{x}))$:

$$\begin{aligned} R(T) &= \int_{\Theta} R_T(\vartheta)\, \mathrm{d}U(\vartheta) \\ &= \int_{\Theta} E_{\vartheta}[L(\vartheta, T)]\, \mathrm{d}U(\vartheta) \\ &= \int_{\Theta}\int_{\mathbb{R}^n} L(\vartheta, T(\mathbf{x}))\, \mathrm{d}F_{\vartheta}(\mathbf{x})\, \mathrm{d}U(\vartheta) \\ &= \int_{\mathbb{R}^n}\int_{\Theta} L(\vartheta, T(\mathbf{x}))\, \mathrm{d}U_{\mathbf{X}}(\vartheta)\, \mathrm{d}F(\mathbf{x}). \end{aligned}$$

Odavde sijedi sljedeće pravilo za konstrukciju Bayesovog procjenitelja:

Teorem 2.1.3. *Za svaku moguću opservaciju $\mathbf{x}$, $\widetilde{T}(\mathbf{x})$ poprima vrijednost koja minimizira*

$$\int_{\Theta} L(\vartheta, T(\mathbf{x}))\, \mathrm{d}U_{\mathbf{X}}(\vartheta).$$

Drugim riječima, za svaku moguću opservaciju $\mathbf{x}$, $\widetilde{T}(\mathbf{x})$ je Bayesov procjenitelj u odnosu na distribuciju $U_{\mathbf{X}}(\vartheta)$.

Napomena
U Bayesovoj statistici, funkciju distribucije $U(\vartheta)$ zovemo *a priori* distribucija od Θ (to je distribucija prije izvršenih promatranja), dok (uvjetnu) funkciju distribucije $U_{\mathbf{X}}(\vartheta)$ zovemo *a posteriori* distribucija od Θ. Takva distribucija sadrži informacije i od slučajnog uzorka i od *a priori* distribucije, ona izražava zaključak statističara o vjerojatnosnoj distribuciji parametra Θ nakon što je dobivena realizacija slučajnog uzorka.

2.2 Bayesova statistika i problem procjene premije

Uveli smo Bayesov način razmišljanja kako bismo formulirali problem procjene premija definiranih u prvom poglavlju. Stoga imamo sljedeća podudaranja:

- funkcional $g(\vartheta)$ podudara se sa točnom individualnom premijom $\mu(\vartheta)$
- Θ kao skup svih mogućih vrijednosti parametara, sa Θ kao skupom svih mogućih individualnih profila rizika ϑ.

Kao i u prvom poglavlju, Θ koristimo također i kao simbol za slučajnu varijablu s funkcijom distribucije $U(\vartheta)$. U kontekstu teorije povjerenja, *a priori* distribuciju $U(\vartheta)$ interpretiramo kao strukturnu distribuciju na skupu Θ.

Za funkciju gubitka odabrat ćemo **funkciju kvadratnog gubitka** (eng. quadratic loss function):

$$L(\vartheta, T(x)) = (\mu(\vartheta) - T(x))^2.$$

Definicija 2.2.1. *Procjenitelj $\widehat{\mu(\vartheta)}$ je dobar barem kao i procjenitelj $\widehat{\widehat{\mu(\vartheta)}}$ (oba za $\mu(\vartheta)$) ako vrijedi*

$$E[(\widehat{\mu(\vartheta)} - \mu(\Theta))^2] \leqslant E[(\widehat{\widehat{\mu(\vartheta)}} - \mu(\Theta))^2].$$

$E[(\widehat{\mu(\vartheta)} - \mu(\Theta))^2]$ zovemo **kvadratni gubitak** (eng. quadratic loss) procjenitelja $\widehat{\mu(\vartheta)}$.

Sljedeći teorem precizno pokazuje što govori definicija 1.3.1, kada kažemo da je $P^{Bayes} = \widetilde{\mu(\vartheta)}$ najbolja moguća premija temeljena na iskustvu.

Teorem 2.2.2. *Bayesov procjenitelj u odnosu na kvadratni gubitak dan je sa*

$$\widetilde{\mu(\vartheta)} = E[\mu(\Theta)|\mathbf{X}]. \tag{2.1}$$

Dokaz. Da bismo pokazali kako je 2.1 zaista Bayesov procjenitelj, dovoljno je pokazati da vrijedi:

$$E[(\widetilde{\mu(\vartheta)} - \mu(\Theta))^2] \leqslant E[(\widehat{\mu(\vartheta)} - \mu(\Theta))^2],$$

za svaki drugi procjenitelj $\widehat{\mu(\vartheta)} = T(\mathbf{X})$.

Neka je $\widehat{\mu(\vartheta)}$ proizvoljan procjenitelj od $\mu(\vartheta)$, a $\widetilde{\mu(\vartheta)}$ neka je *a posteriori* očekivanje $E[\mu(\Theta)|\mathbf{X}]$. Tada vrijedi:

$$\begin{aligned}
E[(\widehat{\mu(\vartheta)} - \mu(\Theta))^2] &= E[E[(\widehat{\mu(\vartheta)} - \mu(\Theta))^2 \,|\mathbf{X}]] \\
&= E[E[(\widehat{\mu(\vartheta)} - \widetilde{\mu(\vartheta)} + \widetilde{\mu(\vartheta)} - \mu(\Theta))^2 \,|\mathbf{X}]] \\
&= E[E[(\widehat{\mu(\vartheta)} - \widetilde{\mu(\vartheta)})^2 \,|\mathbf{X}]] + E[E[(\widetilde{\mu(\vartheta)} - \mu(\Theta))^2 \,|\mathbf{X}]] + \\
&+ 2E[E[(\widehat{\mu(\vartheta)} - \widetilde{\mu(\vartheta)})\,(\widetilde{\mu(\vartheta)} - \mu(\Theta))\,|\mathbf{X}]] \\
&= E[(\widehat{\mu(\vartheta)} - \widetilde{\mu(\vartheta)})^2] + E[(\widetilde{\mu(\vartheta)} - \mu(\Theta))^2] + \\
&+ 2E[(\widehat{\mu(\vartheta)} - \widetilde{\mu(\vartheta)})\,E[(\widetilde{\mu(\vartheta)} - \mu(\Theta))\,|\mathbf{X}]] \\
&= E[(\widehat{\mu(\vartheta)} - \widetilde{\mu(\vartheta)})^2] + E[(\widetilde{\mu(\vartheta)} - \mu(\Theta))^2] \\
&\geq E[(\widetilde{\mu(\vartheta)} - \mu(\Theta))^2]
\end{aligned}$$

U dokazu smo koristili činjenicu da je $\widetilde{\mu(\vartheta)}$ izmjeriva u odnosu na $\mathbf{X}$.
Prema tome slijedi

$$E[(\widetilde{\mu(\vartheta)} - \mu(\Theta))^2] \leqslant E[(\widehat{\mu(\vartheta)} - \mu(\Theta))^2],$$

a odavde slijedi tvrdnja teorema 2.1 □

Teorem 2.2.3. *Vrijede slijedeće dvije relacije:*

(i) Kvadratni gubitak Bayesove premije je

$$E[(\widetilde{\mu(\vartheta)} - \mu(\Theta))^2] = E[Var[\mu(\Theta)|\mathbf{X}]].$$

(ii) Kvadratni gubitak kolektivne premije je

$$\begin{aligned}
E[(\mu_0 - \mu(\Theta))^2] &= Var[\mu(\Theta)] \\
&= E[Var[\mu(\Theta)\,|\mathbf{X}]] + Var[E[\mu(\Theta)\,|\mathbf{X}]].
\end{aligned}$$

Dokaz. (i) Dokaz slijedi direktno iz definicije uvjetne varijance.

$$\begin{aligned}
E[(\widetilde{\mu(\vartheta)} - \mu(\Theta))^2] &= E[E[(\widetilde{\mu(\vartheta)} - \mu(\Theta))^2 \mid \mathbf{X}]] \\
&= E[E[(\mu(\Theta) - E[\mu(\Theta)\,|\mathbf{X}])^2|\mathbf{X}]] \\
&= E[Var[\mu(\Theta)|\mathbf{X}]].
\end{aligned}$$

(ii) Da bismo dokazali ovu relaciju iskoristit ćemo činjenicu da je $\mu_0 = E[\mu(\Theta)]$ i dekompoziciju varijanci:

$$Var[X] = E[Var[X|Z]] + Var[E[X|Z]].$$

Odavde slijedi tvrdnja,

$$\begin{aligned} E[(\mu_0 - \mu(\Theta))^2] &= E[(\mu(\Theta) - E[\mu(\Theta)])^2] \\ &= Var[\mu(\Theta)] \\ &= E[Var[\mu(\Theta)|\mathbf{X}]] + Var[E[\mu(\Theta)|\mathbf{X}]]. \end{aligned}$$

□

Promotrimo sada relacije (i) i (ii). Primjećujemo da imaju zajednički član $E[Var[\mu(\Theta)|\mathbf{X}]]$. Kvadratni gubitak kolektivne premije jednak je kvadratnom gubitku Bayesove premije plus nenegativni član. Dakle, kvadratni gubitak je manji kod Bayesove premije nego kod kolektivne premije. To opravdava činjenicu da je Bayesova premija bolja od kolektivne.

Kroz primjere ćemo kasnije pokazati da to isto vrijedi i za individualnu premiju. Naime, kvadratni gubitak je manji kod Bayesove premije nego kod individualne premije što povlači da je Bayesova premija bolja i od individualne premije, odnosno Bayesova premija je najbolja procjena premije.

2.3 Bayesova premija u tri modela

U svrhu našeg cilja da što točnije odredimo premiju rizika, a budući da smo pokazali da je Bayesova premija najbolja procjena premije, promatrat ćemo Bayesovu premiju u tri specijalna modela:

- model Poisson-gama distribucija
- model binomna-beta distribucija
- model normalna-normalna distribucija

U sva tri modela, prije određivanja Bayesove premije, definirat ćemo:

(i) strukturnu funkciju distribucije $U(\vartheta)$ (*a priori* distribucija od Θ) i

(ii) familiju uvjetnih distribucija od $\mathbf{X}$ uz dano $\Theta = \vartheta$, $\mathcal{F} = \{F_\vartheta : \vartheta \in \Theta\}$.

Prvi dio imena modela predstavlja distribuciju šteta, a drugi dio predstavlja distribuciju parametra. Dakle, cilj nam je procijeniti parametar modela na osnovi opaženih podataka. U biti, na osnovi pretpostavljene, tzv. apriorne distribucije parametara određuje se uvjetna razdioba parametara u odnosu na opaženi uzorak, tzv. aposteriorna razdioba. Na taj način dobit ćemo da je procjena parametra srednja vrijednost aposteriorne razdiobe. Pokušat ću što bolje kombinirati teorijski osvrt i konkretan primjer pojedinog modela.

2.3.1 Poisson-gama model

Kao motivaciju za ovaj model iskoristit ćemo **Bichselov problem**.

Pojavio se problem određivanja visine premije kod automobilskog osiguranja u vrijeme kada se ta premija temeljila isključivo na konjskoj snazi automobila. Naime to nije bila pravedna premija jer su se javljale jako velike razlike među štetama za istu veličinu konjskih snaga. F. Bichsel je bio prvi predstavnik neživotnog osiguranja u Švicarskoj. Bavio se formuliranjem sustava procjene rizika koji bi bio bolje prilagođen individualnom profilu rizika. Pretpostavljao je da bi razlike individualnih profila rizika trebale biti razlika među individualnim brojevima šteta, dok sama veličina šteta, zbog velike varijabilnosti, neće biti dobar pokazatelj očekivane štete.

Uvodimo **matematički model**. Stavimo da nam oznaka N_j predstavlja broj šteta koje učini određeni vozač u godini j, a oznaka X_j predstavlja pripadnu ukupnu štetu vezanu uz tog vozača u godini j.

Bichselov model se temelji na pretpostavci da je

$$E[X_j \mid \Theta = \vartheta] = C \cdot E[N_j \mid \Theta = \vartheta], \ (j = 1, 2, ...). \tag{2.6}$$

Odnosno, očekivani iznos štete u godini j, uz dani individualni profil rizika ϑ, ovisi o konstanti C, koja isključivo ovisi o konjskoj snazi automobila, dok očekivani broj šteta, uz dani $\Theta = \vartheta$, ovisi isključivo o vozaču.

Navedeni model pretpostavlja da frekvencija pojavljivanja šteta ne ovisi o konjskoj snazi automobila, već samo o vozaču kao osobi, međutim u stvarnosti to i nije potpuno točno. Zbog pretpostavke modela 2.6 dovoljno je modelirati broj šteta. Bichselov model za broj šteta temelji se na sljedećim pretpostavkama:

Pretpostavke Poisson-gama modela

- Uvjetno, za dano $\Theta = \vartheta$, N_j, $(j = 1, 2, ...)$ su nezavisne slučajne varijable s Poissonovom distribucijom s parametrom ϑ, tj.

$$P(N_j = k \mid \Theta = \vartheta) = e^{-\vartheta}\,\frac{\vartheta^k}{k!},\ \ k = 0, 1, ... \tag{2.7}$$

- Parametar Θ ima gama distribuciju sa parametrima γ i β, tj. strukturna funkcija ima gustoću:

$$u(\vartheta) = \frac{\beta^\gamma}{\Gamma(\gamma)}\vartheta^{\gamma-1}e^{-\beta\vartheta},\ \ \vartheta > 0. \tag{2.8}$$

Napomena

- Prisjetimo se prva dva momenta gama distribucije slučajne varijable Θ:

$$E[\Theta] = \frac{\gamma}{\beta}\ \ \ i\ \ \ Var[\Theta] = \frac{\gamma}{\beta^2}. \tag{2.9}$$

- Uvodimo oznaku koju koristimo u daljnjem tekstu: $N_\bullet = \sum_{j=1}^n N_j$ gdje točka označava sumiranje po svim pripadnim indeksima.

Iskoristimo pretpostavke Poisson-gama modela 2.7 i 2.8. Uz oznaku F za frekvenciju pojavljivanja štete, vrijedi sljedeća propozicija.

Propozicija 2.3.1. *Vrijede sljedeće relacije:*

$$F^{ind} := E[N_{n+1} \mid \Theta] = \Theta \tag{2.10}$$

$$F^{coll} := E[\Theta] = \frac{\gamma}{\beta} \tag{2.11}$$

$$F^{Bayes} := E[\Theta \mid N_1, N_2, ..., N_n] = \frac{\gamma + N_\bullet}{\beta + n} = \alpha\overline{N} + (1 - \alpha)\frac{\gamma}{\beta}, \tag{2.12}$$

gdje su

$$\alpha = \frac{n}{n + \beta},\ \ \overline{N} = \frac{1}{n}\sum_{j=1}^{n} N_j.$$

Kvadratni gubitak za F_{Bayes} dan je sa:

$$\begin{aligned} E[(F^{Bayes} - \Theta)^2] &= (1 - \alpha)E[(F^{coll} - \Theta)^2] && (2.13)\\ &= \alpha E[(\overline{N} - \Theta)^2] && (2.14) \end{aligned}$$

Prije dokaza propozicije 2.3.1 navest ćemo napomenu:

Napomena

- Veličine premija P_{ind}, P_{coll} i P_{Bayes} dobijemo tako da pripadne veličine F_{ind}, F_{coll} i F_{Bayes} pomnožimo s konstantom C.
- Kvadratni gubitak od P_{Bayes} je

$$E[(P_{Bayes} - P_{ind})^2] = C^2 E[(F_{Bayes} - F_{ind})^2].$$

- Bayesova premija $P_{Bayes} = CF_{Bayes}$ je konveksna funkcija promatranog broja šteta i očekivanje tog broja u kolektivu, pa je time i premija povjerenja (eng. credibility preminum).
- $\alpha = \frac{n}{n+\beta}$ zovemo težina povjerenja (eng. credibility weight):
  - što je veći broj promatranih godina n, težina α će biti veća,
  - što je β veći, težina α će biti manja.

  To nam intuitivno govori da što više informacija imamo o pojedincu, bit će veća važnost koju pridodajemo njegovom iskustvu individualnih šteta, dok s druge strane što je homogeniji kolektiv ($Var[\Theta]$ manja povlači veći β) veću ćemo važnost pridodati iskustvu kolektivne štete za procjenu individualnog rizika, dakle faktor $\frac{\gamma}{\beta}$ u F_{Bayes} će biti veći, tj. težina α će biti manja.
- Iz 2.13 i 2.14 vidimo da vrijedi:

$$\textit{kvadratni gubitak za } F^{Bayes} = (1 - \alpha) \cdot \textit{kvadratni gubitak za } F^{coll},$$
$$\textit{kvadratni gubitak za } F^{Bayes} = \alpha \cdot \textit{kvadratni gubitak za } \overline{N}.$$

  Primijetimo da je procjenitelj F^{coll} baziran samo na *a priori* saznanjima iz kolektiva, a zanemaruje iskustvo individualnih šteta, dok je procjenitelj $\overline{N}$ baziran samo na iskustvu individualnih šteta, a zanemaruje *a priori* saznanja vezana uz kolektiv.
  Iz relacije 2.12 vidimo da procjenitelj F^{Bayes} uzima u obzir oba izvora informacija, te iz relacija 2.13 i 2.14 zbog $\alpha < 1$ vidimo da je kvadratni gubitak za F^{Bayes} manji od kvadratnih gubitaka za F^{coll} i $\overline{N}$.
- Kvadratni gubitak za F^{coll} je

$$E[(F^{coll} - \Theta)^2] = E[(\frac{\gamma}{\beta} - \Theta)^2] = E[(\Theta - E[\Theta])^2] = Var[\Theta] = \frac{\gamma}{\beta^2}. \tag{2.15}$$

Kvadratni gubitak za $\overline{N}$ je

$$E[(\overline{N}-\Theta)^2] = E\{E[(\overline{N}-\Theta)^2|\Theta]\} = E\{Var[\overline{N}|\Theta]\} = E\left[\frac{\Theta}{n}\right] = \frac{E[\Theta]}{n} = \frac{1}{n}\cdot\frac{\gamma}{\beta}. \quad (2.16)$$

Kako je kvadratni gubitak za F^{Bayes} manji od kvadratnog gubitka za $\overline{N}$ vidimo da vrijedi

$$E[(F^{Bayes}-\Theta)^2] \longrightarrow 0 \ \ kada \ \ n \longrightarrow \infty$$

Sada ćemo dokazati propoziciju 2.3.1

Dokaz. Relacija 2.10 slijedi direktno iz pretpostavke da N_{n+1} uz dano Θ ima Poissonovu distribuciju s parametrom Θ.
Budući da je slučajna varijabla Θ gamma distribuirana s parametrima γ i β slijedi relacija 2.11

$$F^{coll} = E[\Theta] = \frac{\gamma}{\beta}. \quad (2.17)$$

Za dokaz tvrdnje 2.12 najprije ćemo odrediti *a posteriori* gustoću od Θ za dano N. Iskoristit ćemo Bayesov teorem:

$$f(x|y) = \frac{f(x,y)}{f(y)} = \frac{f(x)f(y|x)}{\int_{-\infty}^{+\infty} f(x)f(y|x)\mathrm{d}x}. \quad (2.18)$$

Dakle, vrijedi

$$u(\vartheta|N_1,...,N_n) = \frac{\frac{\beta^\gamma}{\Gamma(\gamma)}\vartheta^{\gamma-1}e^{-\beta\vartheta}\prod_{j=1}^n e^{-\vartheta}\frac{\vartheta^{N_j}}{N_j!}}{\int \frac{\beta^\gamma}{\Gamma(\gamma)}\vartheta^{\gamma-1}e^{-\beta\vartheta}\prod_{j=1}^n e^{-\vartheta}\frac{\vartheta^{N_j}}{N_j!}\,\mathrm{d}\vartheta} \propto \vartheta^{\gamma\,+\,\sum_{j=1}^n N_j-1}\ e^{-(\beta+n)\ \vartheta}, \quad (2.19)$$

gdje je $\propto$ znak proporcionalnosti, tj. desna strana jednaka je lijevoj strani pomnoženoj s konstantom koja ne ovisi o ϑ.
Iz gornjeg izraza zaključujemo da je *a posteriori* distribucija od Θ za dano $(N_1,...,N_n)$ gama distribucija s parametrima:

$$\gamma' = \gamma + N_\bullet,$$
$$\beta' = \beta + n.$$

Odavde slijedi prva jednakost relacije 2.12

$$F^{Bayes} = E[\Theta|N_1,\ldots,N_n] = \frac{\gamma'}{\beta'} = \frac{\gamma+N_\bullet}{\beta+n}.$$

Dalje imamo

$$\begin{aligned}
\frac{\gamma + N_\bullet}{\beta + n} &= \frac{1}{\beta + n} \cdot N_\bullet \ + \ \frac{1}{\beta + n} \cdot \gamma \\
&= \frac{n}{\beta + n} \cdot \frac{1}{n} \cdot N_\bullet \ + \ \frac{\beta}{\beta + n} \cdot \frac{1}{\beta} \cdot \gamma \\
&= \frac{n}{\beta + n} \cdot \frac{1}{n} \cdot n \cdot \overline{N} \ + \ \left(1 \ - \ \frac{n}{n + \beta}\right) \cdot \frac{1}{\beta} \cdot \gamma \\
&= \alpha \overline{N} \ + \ (1 \ - \ \alpha)\frac{\gamma}{\beta},
\end{aligned}$$

gdje je

$$\alpha = \frac{n}{n \ + \ \beta}, \quad \overline{N} = \frac{1}{n} \sum_{j=1}^{n} N_j.$$

Time smo dokazali i drugu nejednakost relacije 2.12.
Da bismo dokazali relacije 2.13 i 2.14 za kvadratni gubitak od F^{Bayes} koristiti ćemo teorem 2.2.3 i sljedeće distribucije:

- $\Theta \sim \Gamma(\gamma, \beta)$
- uvjetno, za dano $(N_1, ..., N_n)$, Θ je gama distribuirana s parametrima $\gamma' = \gamma + N_\bullet$ i $\beta' = \beta + n$
- za dano $\Theta = \vartheta$, N_j, $\ j = 1, 2, ..., n$ imaju Poissonovu distribuciju s parametrom ϑ.

Sada imamo

$$\begin{aligned}
E[(F^{Bayes} - \Theta)^2] = E[Var[\Theta|N]] = E\left[\frac{\gamma + N_\bullet}{(\beta + n)^2}\right] &= \frac{1}{(\beta + n)^2} E[\gamma + N_\bullet] \\
&= \frac{\gamma}{(\beta + n)^2}\left(1 + \frac{E[N_\bullet]}{\gamma}\right) = \frac{\gamma}{(\beta + n)^2}\left(1 + \frac{E[E[N_\bullet|\Theta]]}{\gamma}\right) \\
&= \frac{\gamma}{(\beta + n)^2}\left(1 + \frac{E\left[\sum_{j=1}^n E[N_j|\Theta]\right]}{\gamma}\right) = \frac{\gamma}{(\beta + n)^2}\left(1 + \frac{E[n \cdot \Theta]}{\gamma}\right) \\
&= \frac{\gamma}{(\beta + n)^2}\left(1 + \frac{n \cdot \frac{\gamma}{\beta}}{\gamma}\right) = \frac{\gamma}{(\beta + n)^2}\left(1 + \frac{n}{\beta}\right) \\
&= \frac{\gamma}{(\beta + n)^2} \cdot \frac{n + \beta}{\beta} = \frac{\gamma}{\beta} \cdot \frac{1}{n + \beta} = \frac{\gamma}{\beta^2} \cdot \frac{\beta}{n + \beta} = \frac{\gamma}{\beta^2} \cdot \left(1 \ - \ \frac{n}{n + \beta}\right) \\
&= \frac{\gamma}{\beta^2} \cdot (1 \ - \ \alpha) = (1 \ - \ \alpha) \cdot E[(F^{cool} \ - \ \Theta)^2].
\end{aligned}$$

Time smo dokazali relaciju 2.13.
Iz

$$E[(F^{Bayes} - \Theta)^2] = \frac{\gamma}{\beta} \cdot \frac{1}{n+\beta} = \frac{\gamma}{\beta} \cdot \frac{n}{n+\beta} \cdot \frac{1}{n} = \frac{1}{n} \cdot \frac{\gamma}{\beta} \cdot \alpha = \alpha \cdot E[(\overline{N} - \Theta)^2],$$

vrijedi relacija 2.13.

□

Sada želimo procijeniti nepoznate strukturne parametre γ i β.
Prema pretpostavci Poisson-gama modela 2.8 znamo distribuciju slučajne varijable $\Theta \sim \Gamma(\gamma, \beta)$. Dakle, da bismo procijenili Θ trebamo procijeniti nepoznate parametre γ i β.
Za prikaz problema koristit ćemo podatke o broju šteta kod švicarskih vozača automobila iz 1961 godine, koji su bili poznati F. Bichselu.

k	0	1	2	3	4	5	6	ukupno
broj polica s k šteta	103 704	14 075	1766	255	45	6	2	119 853

Tablica 2.1: Podaci o švicarskim vozačima automobila

Tablicu možemo čitati na ovaj način:

Imamo 119 853 ponavljana eksperimenta "dvije urne" pa je za svaki od 119 853 vozača individualni profil rizika izabran iz prve urne, a pripadni broj šteta izabran je iz druge urne s tim da biranje iz druge urne ovisi o već odabranom individualnom profilu rizika iz prve urne.

Dakle, imamo skup parova $\{ (\Theta_1, N^{(1)}), \dots, (\Theta_i, N^{(i)}), \dots, (\Theta_I, N^{(I)})\}$, gdje je $I = 119\,853$, i označava pojedinog vozača, a $N^{(i)}$ je njegov pripadni broj šteta u 1961 godini.

Postoje brojni načini kako možemo na temelju podataka tablice 2.1 procijeniti strukturne parametre γ i β. Bichsel je koristio tzv. metodu momenata. Ona kaže sljedeće:
Uz pretpostavku da su parovi slučajnih varijabli $\{(\Theta_i, N^{(i)}) \ : \ i = 1, 2, ..., I\}$ nezavisni i jedanko distribuirani, sljedeći empirijski momenti su nepristrani procjenitelji:

$$\widehat{\mu}_N = \overline{N}_{1961} = \frac{1}{I} \sum_{i=1}^{I} N^{(i)} = 0.155$$

$$\widehat{\sigma}_N^2 = \frac{1}{I-1} \sum_{i=1}^{I} \left(N^{(i)} - \overline{N}_{1961}\right)^2 = 0.179.$$

S druge, odredimo prva dva momenta slučajne varijable N, pri čemu prema pretpostavkama modela znamo da je slučajna varijabla Θ gamma distribuirana s parametrima γ i β, a da je uvjetna distribucija od N_j, uz dano $\Theta = \vartheta$, Poissonova s parametrom ϑ. Još ćemo napomenuti da su očekivanje i varijanca Poissonove slučajne varijable s parametrom Θ, jednaki upravo tom parametru Θ.

Odavde slijedi:

$$\mu_N = E[N] = E[E[N|\Theta]] = E[\Theta] = \frac{\gamma}{\beta},$$
$$\sigma_N^2 = Var[N] = E[Var[N|\Theta]] + Var[E[N|\Theta]] = E[\theta] + Var[\Theta] = \frac{\gamma}{\beta} \cdot (1 + \frac{1}{\beta}).$$

Da bi dobili procjene parametara, izjednačimo:

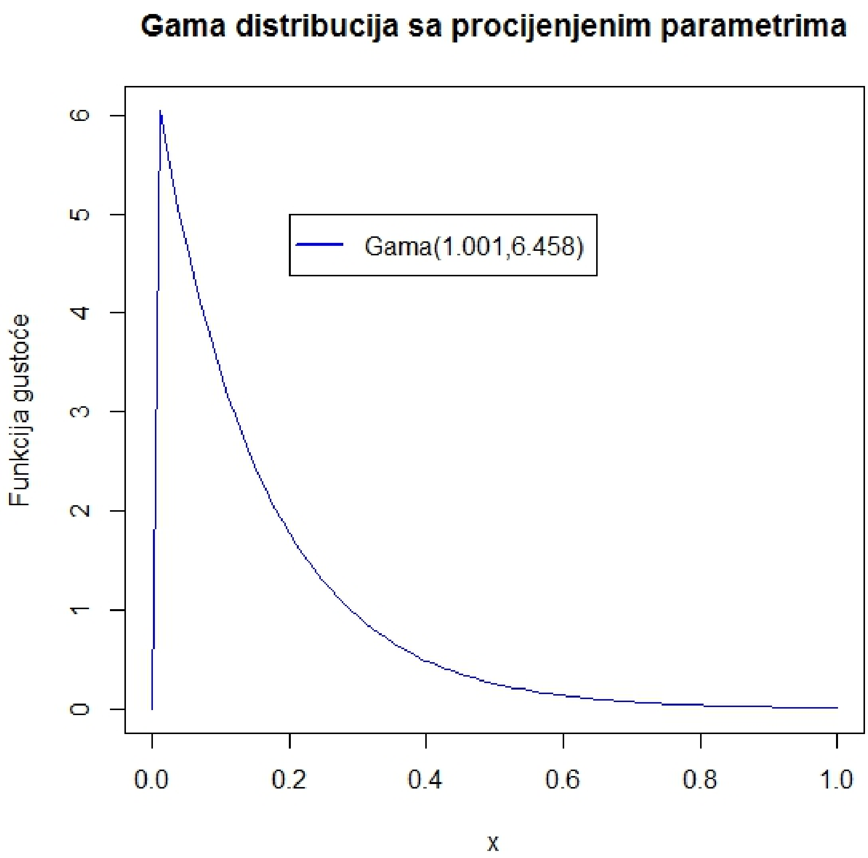
$$\widehat{\mu}_N = \mu_N$$
$$\widehat{\sigma}_N^2 = \sigma_N^2.$$

Stoga za naš primjer imamo:

$$\frac{\widehat{\gamma}}{\widehat{\beta}} = 0.155$$
$$\frac{\widehat{\gamma}}{\widehat{\beta}} \cdot (1 + \frac{1}{\widehat{\beta}}) = 0.155 \cdot (1 + \frac{1}{\widehat{\beta}}) = 0.179.$$

A odavde rješavanjem sustava dobijemo procjenu za strukturne parametre γ i β, tj.

$$\widehat{\gamma} = 1.001$$
$$\widehat{\beta} = 6.458.$$



Slika 2.2: Grafički prikaz funkcije gustoće s procijenjenim strukturnim parametrima.

Ako sada umjesto nepoznatih strukturnih parametara γ i β u formulu za Bayesov procjenitelj 2.12 uvrstimo procijenjene parametre $\widehat{\gamma}$ i $\widehat{\beta}$ iz kolektiva, dobit ćemo **empirijski Bayesov procjenitelj**:

$$\widehat{F^{Bayes}}^{emp} = \frac{n}{n+\widehat{\beta}}\overline{N} + \left(1 - \frac{n}{n+\widehat{\beta}}\right)\frac{\widehat{\gamma}}{\widehat{\beta}} = \frac{n}{n+\widehat{\beta}}\overline{N} + \frac{\widehat{\beta}}{n+\widehat{\beta}} \cdot \frac{\widehat{\gamma}}{\widehat{\beta}},$$

gdje je

$$\overline{N} = \frac{1}{n}\sum_{j=1}^{n} N_j.$$

Dakle, da bismo odredili Bayesovu premiju koristimo sljedeća dva koraka:

- određivanje Bayesove premije P^{Bayes} pod pretpostavkom da su strukturni parametri poznati,
- procjena strukturnih parametara koristeći podatke iz kolektiva.

Distribucija rizika u kolektivu
Koliko dobro ovaj model odgovara podacima iz tablice o švicarskim vozačima (2.1) provjerit ćemo tako da prvo nađemo bezuvjetnu distribuciju od N.

Za gornji slučaj, eksplicitno ćemo odrediti distribuciju:

$$\begin{aligned}
P(N=k) &= \int P(N=k|\Theta=\vartheta)u(\vartheta)\mathrm{d}\vartheta = \int e^{-\vartheta}\frac{\vartheta^k}{k!}\frac{\beta^\gamma}{\Gamma(\gamma)}\vartheta^{\gamma-1}e^{-\beta\vartheta}\mathrm{d}\vartheta \\
&= \frac{\beta^\gamma}{\Gamma(\gamma)}\cdot\frac{1}{k!}\underbrace{\int e^{-(\beta+1)\vartheta}\vartheta^{\gamma+k-1}\mathrm{d}\vartheta}_{=\frac{\Gamma(\gamma+k)}{(\beta+1)^{\gamma+k}}} = \frac{\Gamma(\gamma+k)}{\Gamma(\gamma)k!}\left(\frac{\beta}{\beta+1}\right)^\gamma\left(\frac{1}{\beta+1}\right)^k \\
&= \binom{\gamma+k-1}{k}p^\gamma(1-p)^k, \quad sa \;\; p=\frac{\beta}{\beta+1}
\end{aligned}$$

a to je negativna binomna distribucija.

Promatrat ćemo dva slučaja:

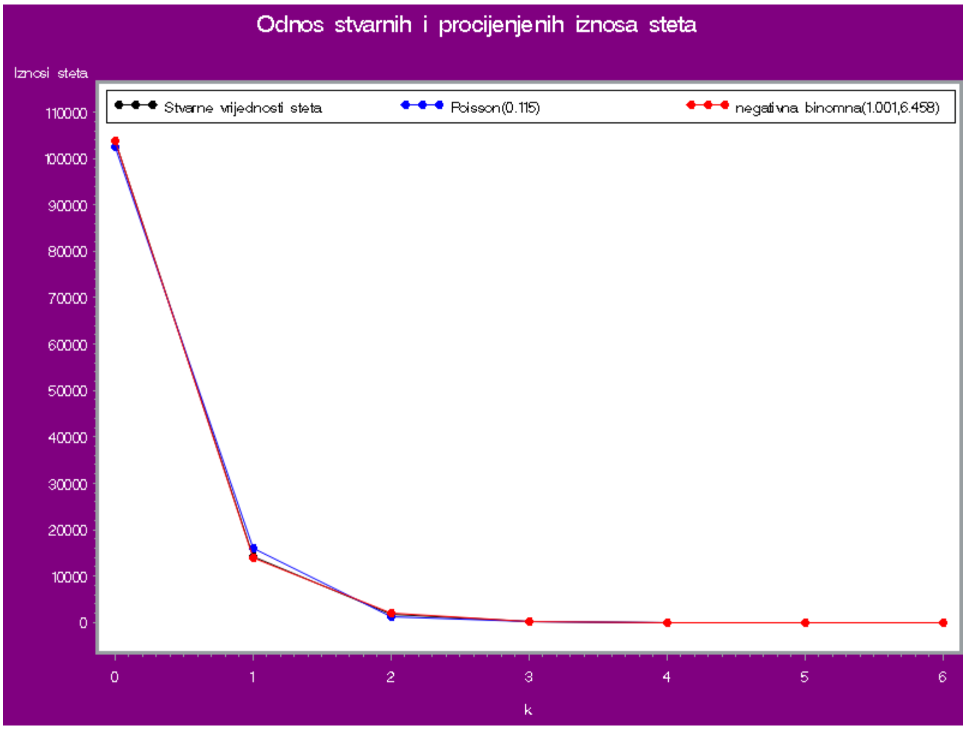
- slučaj homogenog portfelja u kojem sve police imaju Poissonovu distribuciju s istim parametrom λ,
- model ”dvije urne” uz Poisson-gama pretpostavku što povlači da je bezuvjetna distribucija od N negativna binomna distribucija s parametrima γ i β.

U oba slučaja ćemo izračunati očekivani broj polica s $k = 0, 1, ..., 6$ šteta te dobivene vrijednosti usporediti sa stvarnim brojem polica vezanih uz isti broj šteta. Na taj način vidjeti ćemo koja distribucija daje bolje rezultate.

k	stvarna vrijednost	Poisson($\lambda = 0.155$)	neg. binomna($\gamma = 1.001, \beta = 6.458$)
0	103 704	102 629	103 757
1	14 075	15 922	13 934
2	1 766	1 234	1 871
3	255	64	251
4	45	3	34
5	6	0	5
6	2	0	1
ukupno	119 853	119 853	119 853

Tablica 2.2: Broj polica sa k šteta u tri slučaja

Grafički prikaz rezultata:



Slika 2.3: Odnos između stvarnog i očekivanog broja polica s $k = 0, 1, ..., 6$ šteta.

Rezultati iz tablice 2.2 i slika 2.3 pokazuju da negativna binomna distribucija daje pouzdanije rezultate.

Primjena

Ovaj model i podaci poslužili su Bichselu kao osnova za konstrukciju bonus-malus sustava u obveznom automobilskom osiguranju u Švicarskoj. Njegov cilj je bio prilagoditi individualnu premiju prema procijenjenoj očekivanoj frekvenciji šteta pojedinog vozača. Ideja određivanja individualne premije je:

- ako za dani rizik očekujemo dvostruko veći broj šteta u odnosu na prosječnog vozača, onda je fer da njegova premija bude dvostruko veća od premije koju plaća prosječni vozač
- ako za vozača očekujemo da će napraviti dvostruko manji broj šteta od prosjeka, onda će njegova premija biti dvostruko manja od prosječne premije, odnosno taj će vozač dobiti bonus od 50%.

Dakle, da bismo dobili premiju za pojedinog vozača, moramo kolektivnu premiju pomnožiti s bonus-malus faktorom:

$$bonus - malus \ \ faktor = \frac{\widehat{F}^{Bayes}}{\lambda_0},$$

gdje je λ_0 prosječna frekvencija šteta, ona kod nas iznosi $\lambda_0 = 0.155$. Bonus-malus faktor može poprimiti vrijednost i manju i veću od jedan. Ukoliko je manji od jedan, premija za tog vozača bit će manja od kolektivne premije i obratno.
U donjoj tablici dane su vrijednosti bonus-malus faktora određenih za naš primjer.

$n \setminus k$	0	1	2	3
1	87%	173%	260%	346%
2	76%	153%	229%	305%
3	68%	136%	205%	273%
4	62%	123%	185%	247%
5	56%	113%	169%	225%
6	52%	104%	155%	207%

Tablica 2.3: Bonus-malus faktor izražen u postocima, gdje je n promatrani broj godina, a k je promatrani broj šteta

Tablicu 2.3 možemo tumačiti na sljedeći način: 52% u lijevom donjem kutu znači: ukoliko vozač u 6 godina ne napravi niti jednu štetu, njegov će bonus-malus faktor iznositi 52%

prosječne premije, dakle njegova premija će biti gotovo dvostruko manja od prosječne premije. S druge strane, ukoliko vozač u 6 godina napravi 3 štete, platit će premiju i više nego dvostruko veću u odnosu na prosječnu premiju. Njegov pripadni bonus-malus faktor iznosi 207% prosječne premije.

Model u kojem se *a priori* očekivana frekvencija šteta razlikuje među godinama i rizicima

Navest ćemo dva primjera koji pobijaju Bichselovu pretpostavku. Naime on je u svom modelu pretpostavio da svi vozači imaju *a priori* jednake rizike, tj. da *a priori* svi vozači imaju jednaku očekivanu frekvenciju šteta. Međutim, u mnogim situacijama u praksi to nije slučaj.

Na primjer, promatramo li rizik za osiguranu grupu automobila, očekivana frekvencija šteta svake grupe neće biti *a priori* jednaka, već će ovisiti o tipovima automobila u toj grupi, o starosti pojedinih automobila, broju prijeđenih kilometara i sličnim eksplanatornim varijablama.

Tako na primjer, ako promatramo rizik kod grupnog životnog osiguranja ili grupnog osiguranja od nezgode, očekivani iznos premije će ovisiti o činjenicama kao što su starost ili spolna struktura grupe. Također, mogu biti prisutne promjene ili trendovi u frekvenciji šteta tokom godine, ali u svakom slučaju, ne možemo uzeti da su *a priori* svi rizici jednaki.

Sljedeće pretpostavke su ekvivalentne pretpostavkama 2.7 i 2.8:

- Uvjetno, za dano $\Theta = \vartheta$, N_j, $(j = 1, 2, ...)$ su nezavisne slučajne varijable s Poissonovom distribucijom s parametrom $\vartheta \cdot \lambda_0$, gdje je $\lambda_0 = \frac{\gamma}{\beta}$ *a priori* očekivana frekvencija šteta.
- Θ ima gama distribuciju sa $E[\Theta] = 1$ i parametrom oblika γ.

Napomena

- Primijetimo da iz uvjeta $E[\Theta] = 1$ slijedi da su parametri gama distribucije slučajne varijable Θ jednaki, tj, $\gamma = \beta$. Naime,
$$E[\Theta] = \frac{\gamma}{\beta} \text{ i } E[\Theta] = 1 \;\; => \;\; \frac{\gamma}{\beta} = 1 \;\; => \;\; \gamma = \beta.$$
- Varijanca slučajne varijable Θ je: $Var[\Theta] = \frac{\gamma}{\beta^2} = \frac{\gamma}{\gamma^2} = \frac{1}{\gamma}$.
- Kako je uvjetna distribucija od N_j uz dano Θ (kraće, u oznaci $(N_j|\Theta)$) Poissonova slučajna varijabla, vrijedi: $F^{ind} = E[N_j|\Theta] = \lambda_0 \cdot \Theta$.

- Dobra intuitivna mjera heterogenosti portfelja je koeficijent varijacije $CoVa(F^{ind})$. U ovom modelu je $CoVa(F^{ind}) = \sqrt{Var[\Theta]} = \sqrt{\frac{1}{\gamma}} = \frac{1}{\sqrt{\gamma}}$ i parametar γ sada ima direktnu interpretaciju $\gamma = (CoVa(F^{ind}))^{-2}$.
- Za procjenu Θ, umjesto apsolutne frekvencije šteta N_j, promatramo relativnu frekvenciju šteta $\widetilde{F}_j = \frac{N_j}{\lambda_0}$.
- Kako je $(N_j|\Theta = \vartheta) \sim P(\vartheta \cdot \lambda_0)$, slijedi:
$$\widetilde{F}^{ind} = E[\widetilde{F}_j|\Theta] = E\left[\frac{N_j}{\lambda_0}|\Theta\right] = \frac{1}{\lambda_0} \cdot E[N_j|\Theta] = \frac{1}{\lambda_0} \cdot \lambda_0 \cdot \Theta = \Theta.$$
- Izvod formule za $\widetilde{F}^{Bayes}$:
$$\begin{aligned}\widetilde{F}^{Bayes} &= \frac{F^{Bayes}}{\lambda_0} = \frac{\alpha \cdot \overline{N} + (1-\alpha) \cdot \frac{\gamma}{\beta}}{\lambda_0} = \frac{\frac{n}{n+\beta} \cdot \frac{1}{n} \cdot N_{\bullet} + \frac{\beta}{n+\beta} \cdot \frac{\gamma}{\beta}}{\lambda_0} \\ &= \frac{N_{\bullet}}{\lambda_0 \cdot (n+\beta)} + \frac{\gamma}{\lambda_0 \cdot (n+\beta)} = 1 + \frac{N_{\bullet}}{\lambda_0 \cdot (n+\beta)} + \frac{\gamma}{\lambda_0 \cdot (n+\beta)} - 1 \\ &= 1 + \frac{n}{n+\beta}\left(\frac{N_{\bullet}}{n\lambda_0} + \frac{\gamma}{n\lambda_0} - \frac{(n+\beta)\lambda_0}{n\lambda_0}\right) = 1 + \alpha\frac{N_{\bullet} + \gamma - \beta\lambda_0 - n\lambda_0}{n\lambda_0} \\ &= 1 + \alpha\left(\frac{N_{\bullet} + (\gamma - \beta\lambda_0)}{n\lambda_0} - 1\right) = 1 + \alpha\left(\frac{N_{\bullet}}{n\lambda_0} - 1\right) = 1 + \alpha\left(\frac{N_{\bullet}}{\nu_{\bullet}} - 1\right),\end{aligned}$$
gdje je
$$\nu_{\bullet} = n\lambda_0 \text{ } \textit{a priori} \text{ očekivani broj šteta.}$$
Također vrijedi: $\overline{\widetilde{F}} = \frac{N_{\bullet}}{\nu_{\bullet}} = \frac{\sum_{j=1}^n N_j}{n\lambda_0} = \frac{1}{n}\sum_{j=1}^n \frac{N_j}{\lambda_0} = \frac{1}{n}\sum_{j=1}^n \widetilde{F}_j$,

a odavde sijedi $\widetilde{F}^{Bayes} = \widetilde{\Theta} = 1 + \alpha \cdot (\overline{\widetilde{F}} - 1)$.

Sada promatramo situaciju kada *a priori* očekivana frekvencija šteta varira među godinama i među rizicima u portfelju. Za određeni rizik sa N_j označit ćemo broj šteta u godini j, a sa λ_j *a priori* očekivani broj šteta u toj godini za taj rizik. Naravno, λ_j može ovisiti o određenim karakteristikama kao što su dobna i spolna struktura osigurane grupe.
Formirat ćemo sljedeću pretpostavku koja zadovoljava naš problem:

Pretpostavke Poisson-gama modela (model II)

- Uvjetno, za dano Θ, slučajne varijable N_j, $(j = 1, 2, ...)$ su nezavisne s Poissonovom distribucijom s parametrom $\Theta \cdot \lambda_j$, gdje je λ_j *a priori* očekivani broj šteta u godini j.
- Slučajna varijabla Θ ima gama distribuciju sa $E[\Theta] = 1$ i parametrom oblika γ.

Propozicija 2.3.2. *Uz upravo navedenu pretpostavku vrijedi:*

$$\widetilde{F}^{Bayes} = \widetilde{\Theta} = E[\Theta|N_1, \dots, N_n] = 1 + \alpha\left(\frac{N_\bullet}{\nu_\bullet} - 1\right), \tag{2.20}$$

gdje je

$$\begin{aligned}
N_\bullet &= \sum_{j=1}^{n} N_j = \textit{promatrani broj šteta},\\
\nu_\bullet &= \sum_{j=1}^{n} \lambda_j = \textit{a priori očekivani broj šteta},\\
\alpha &= \frac{\nu_\bullet}{\nu_\bullet + \gamma}.
\end{aligned}$$

Formula 2.20 je vrlo praktična i intuitivna. Ona kaže da je $\widetilde{F}^{Bayes}$ jednak jedan plus pripadna korekcija za promatrani rizik, a ta korekcija predstavlja težinu povjerenja α pomnoženu sa odstupanjem odnosa ostvarenog broja šteta i pripadnog očekivanog broja šteta.

Primjer 2.3.3 (Numerička ilustracija Poisson-gama modela)**.** *Prepostavimo da sve štete u portfelju iznose točno 1. Tako da je ukupna šteta u nekom periodu određena jedino njihovim brojem. Prepostavimo nadalje da broj šteta u svakoj godini ima Poissonovu razdiobu s parametrom ϑ. Prepostavimo da parametar ϑ nije poznat, no na osnovu prethodnog iskustva i drugih podataka utvrđeno da je ϑ približno ima $\Gamma(\gamma, \beta)$ razdiobu.*

Pretpostavimo da u 10 sukcesivnih godina broj šteta ima sljedeće vrijednosti:

$$144,\ 144,\ 174,\ 148,\ 151,\ 156,\ 168,\ 147,\ 140,\ 161$$

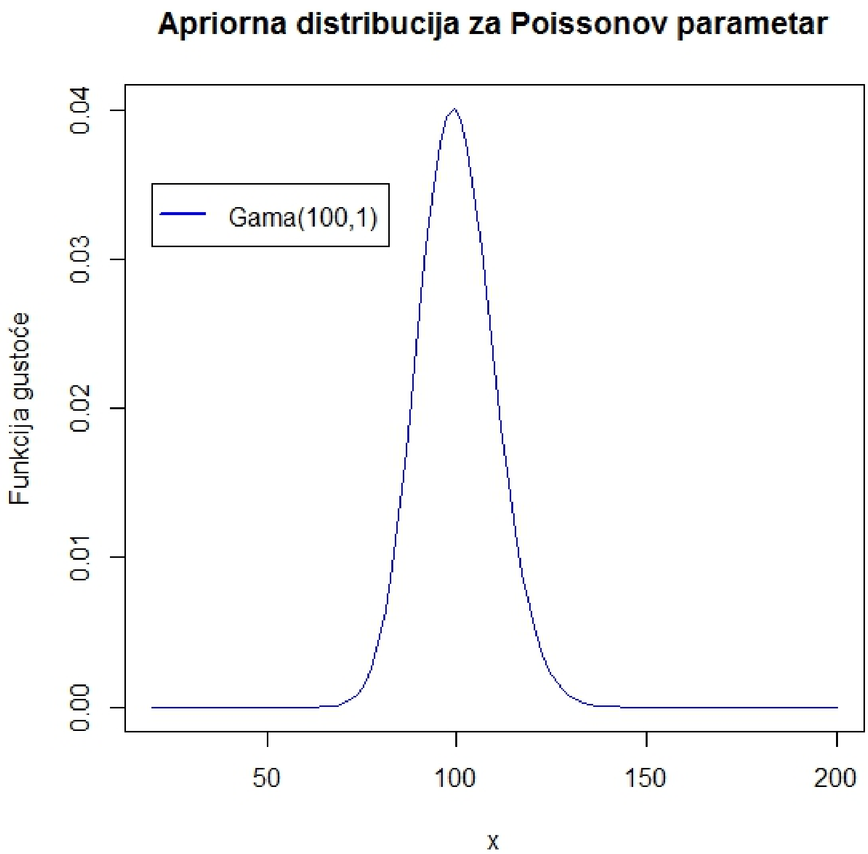
Ovi brojevi dolaze iz Poisson-gama modela uz (brojevi su simulirani pa znamo da dolaze iz $P(\vartheta)$, $\vartheta = 150$) dvije a priori razdiobe za parametar ϑ

$$\Gamma(100, 1)\ \ i\ \ \Gamma(500, 5).$$

Tada prema formuli a posteriori razdiobe od parametra ϑ $\Gamma(k_1 + ... + k_n + \gamma, \beta + n)$ *imamo sljedeće a posteriori razdiobe:*

$$\Gamma(1633, 11) \ \ i \ \ \Gamma(2033, 15).$$

Zbog prevelikih iznosa parametara gamma distribucije, grafički možemo prikazati samo a priori razdiobu $\Gamma(100, 1)$*.*



Želimo odrediti premiju na osnovu ovih pretpostavki. Bayesovski procjenitelj za premiju dan je formulom

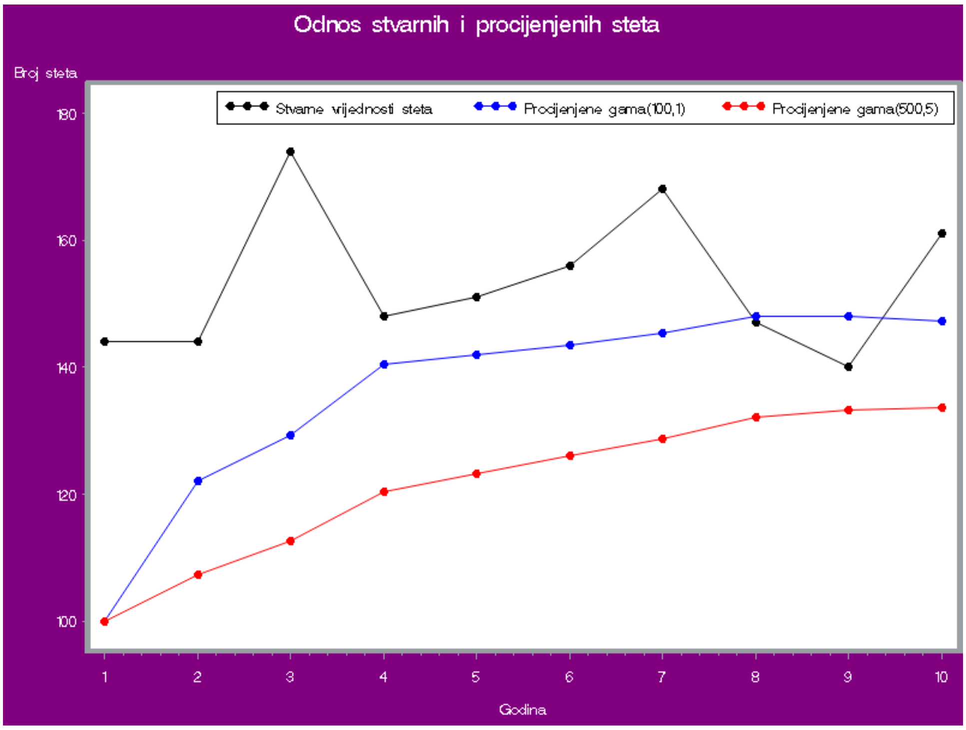
$$E[\Theta|N_1, N_2, \ldots, N_n] = \frac{N_1 + N_2 + \cdots + N_n + \gamma}{n + \beta}.$$

Primjetimo da je apriori očekivanje od parametra ϑ *(pa dakle i od* N_j, $j = 1, 2, ..., 10$*) u oba slučaja jednako* $\frac{\gamma}{\beta} = 100$*, dok je varijanca kod* $\Gamma(100, 1)$ *jednaka 100, a kod* $\Gamma(500, 5)$ *jednaka 20.*

Rezulate procijenjenih broj šteta u ovisnosti distribucije za parametar ϑ koje sam dobila računom prikazat ću u sljedećoj tablici:

godina	stvarne vrijednosti šteta	procijenjene za $\Gamma(100, 1)$	procijenjene za $\Gamma(500, 5)$
1	144	100	100
2	144	122.0	107.3
3	174	129.3	112.6
4	148	140.5	120.3
5	151	142.0	123.3
6	156	143.5	126.1
7	168	145.3	128.8
8	147	148.1	132.1
9	140	148.0	133.2
10	161	147.2	133.7

Grafički prikaz rezultata:



Slika 2.4: Odnosi između stvarnih vrijednosti šteta i procijenjenih po godinama

Slika 2.4 pokazuje da iako obje apriori razdiobe imaju očekivanje 100, razdioba $\Gamma(500, 5)$ *ima bitno manju varijancu, što uzrokuje sporiju konvergenciju Bayesovskog procjenitelja k stvarnoj vrijednosti.*

2.3.2 Binomni-beta model

Motivacija

U grupnom životnom osiguranju ili u grupnom osiguranju od nezgoda, između ostalog, zanima nas broj šteta, odnosno frekvencija pojavljivanja šteta u grupi. Radi jednostavnosti pretpostavljamo da je uz svakog člana grupe vezana jednaka vjerojatnost štete, da se individualne štete javljaju nezavisno te da članovi koji prouzroče štetu napuštaju grupu.

Štetu u ovom kontekstu možemo protumačiti kao invaliditet ili smrt osigurane osobe što predstavlja štetu za osiguravajuću kuću. U tom slučaju osiguranik napušta grupu, a osiguravajuća kuća mu je dužna isplatiti naknadu.

Za svaku godinu $j = 1, 2, ...$ definirat ćemo slijedeće slučajne varijable:

$$
\begin{aligned}
N_j &= \text{ broj šteta koje se javljaju u promatranoj grupi u godini } j,\\
V_j &= \text{broj članova u grupi na početku godine } j,\\
X_j &:= \frac{N_j}{V_j} = \text{ frekvencija šteta u godini } j.
\end{aligned}
$$

U svakom trenutku n, sve slučajne varijable s indeksom manjim ili jednakim n su poznate, a mi želimo odrediti frekvenciju šteta za godinu $n+1$:

$$X_{n+1} = \frac{N_{n+1}}{V_{n+1}}.$$

Primjećujemo da je na desnoj strani gornje jednadžbe V_{n+1} poznata, odnosno jedino je još N_{n+1} nepoznata. Stoga želimo modelirati N_{n+1}. Taj će se model temeljiti na sljedećoj pretpostavci:

Pretpostavke binomnog-beta modela

- uvjetno, za dano $\Theta = \vartheta$, slučajne varijable N_j, $(j = 1, 2, ...)$ su nezavisne s binomnom distribucijom, tj. vrijedi:

$$P[N_j = k|\Theta = \vartheta] = \binom{V_j}{k} \vartheta^k (1-\vartheta)^{V_j - k}. \tag{2.21}$$

- slučajna varijabla Θ ima beta $B(a, b)$ distribuciju sa parametrima $a, b > 0$, što je ekvivalentno s tim da strukturna funkcija ima gustoću:

$$u(\vartheta) = \frac{1}{B(a,b)}\vartheta^{a-1}(1-\vartheta)^{b-1},\ 0 \leq \vartheta \leq 1,\ \text{gdje je } B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}. \tag{2.22}$$

Prva dva momenta beta distribuirane slučajne varijable Θ su:

$$E[\Theta] = \frac{a}{a+b} \quad i \quad Var[\Theta] = \frac{ab}{(1+a+b)(a+b)^2}.$$

Propozicija 2.3.4. *Uz pretpostavku binomnog-beta modela, vrijede sljedeće relacije:*

$$F^{ind} = E[X_{n+1} \mid \Theta] = \Theta \tag{2.23}$$

$$F^{coll} = E[\Theta] = \frac{a}{a\ +\ b} \tag{2.24}$$

$$F^{Bayes} = \frac{a\ +\ N_\bullet}{a\ +\ b +\ V_\bullet} = \alpha\overline{N} + (1-\alpha)\frac{a}{a\ +\ b},\ \textit{gdje je}\ \overline{N} = \frac{N_\bullet}{V_\bullet},\ \alpha = \frac{V_\bullet}{a\ +\ b\ +\ V_\bullet}. \tag{2.25}$$

Kvadratni gubitak za F^{Bayes} dan je sa:

$$\begin{aligned} E[(F^{Bayes}\ -\ \Theta)^2] &= (1\ -\ \alpha)E[(F^{coll}\ -\ \Theta)^2] \quad (2.26)\\ &= \alpha E[(\overline{N}\ -\ \Theta)^2]. \quad (2.27)\end{aligned}$$

Dokaz. Kako je uvjetna distribucija od N_{n+1} za dano Θ binomna s parametrima V_{n+1} i Θ, vrijedi:

$$F^{ind} = E[X_{n+1}|\Theta] = E\left[\frac{N_{n+1}}{V_{n+1}}\right] = \frac{1}{V_{n+1}}E[N_{n+1}|\Theta] = \frac{1}{V_{n+1}}V_{n+1}\Theta = \Theta.$$

Budući da Θ ima beta distribuciju s parametrima a i b, direktno slijedi:

$$F^{coll} = E[\Theta] = \frac{a}{a+b}.$$

Da bismo dobili F^{Bayes}, najprije ćemo odrediti *a posteriori* gustoću od slučajne varijable Θ za dane $N_1, N_2, ..., N_n$. Pri tome koristimo Bayesov teorem:

$$u(\vartheta|N_1 = k_1, ..., N_n = k_n) = \frac{\frac{1}{B(a,b)} \cdot \vartheta^{a-1} \cdot (1-\vartheta)^{b-1} \cdot \prod_{j=1}^{n} \binom{V_j}{k_j} \vartheta^{k_j}(1-\vartheta)^{V_j-k_j}}{\int_0^1 \frac{1}{B(a,b)} \cdot \xi^{a-1} \cdot (1-\xi)^{b-1} \cdot \prod_{j=1}^{n} \binom{V_j}{k_j} \xi^{k_j}(1-\xi)^{V_j-k_j} \mathrm{d}\xi},$$

a odavde slijedi

$$u(\vartheta|N_1,...,N_n) \;\propto\; \prod_{j=1}^{n} \begin{pmatrix} V_j \\ N_j \end{pmatrix} \vartheta^{N_j}(1-\vartheta)^{V_j-N_j}\vartheta^{a-1}(1-\vartheta)^{b-1} \;\propto\; \vartheta^{a+N_\bullet-1}(1-\vartheta)^{b+V_\bullet-N_\bullet-1}.$$

Dakle, *a posteriori* distribucija slučajne varijable Θ ponovno je beta distribucija, ali sada sa novim parametrima

$$a' = a + N_\bullet \quad \text{i} \quad b' = b + V_\bullet - N_\bullet.$$

Iz upravo dokazanoga slijedi da je

$$F^{Bayes} \;=\; \widetilde{\Theta} \;=\; E[\Theta|N_1,...,N_n] \;=\; \frac{a'}{a'+b'} \;=\; \frac{a+N_\bullet}{a+N_\bullet+b+V_\bullet-N_\bullet} \;=\; \frac{a+N_\bullet}{a+b+V_\bullet},$$

što smo i trebali dokazati.

Sada ćemo dokazati i drugu jednakost u izrazu za F^{Bayes}.

$$\begin{aligned} F^{Bayes} &= \frac{a+N_\bullet}{a+b+V_\bullet} = \frac{N_\bullet}{a+b+V_\bullet} + \frac{a}{a+b+V_\bullet} \\ &= \frac{V_\bullet}{a+b+V_\bullet}\cdot\frac{N_\bullet}{V_\bullet} + \frac{a+b}{a+b+V_\bullet}\cdot\frac{a}{a+b} \\ &= \alpha\overline{N} \;+\; (1-\alpha)\frac{a}{a+b}, \end{aligned}$$

gdje je

$$\overline{N} = \frac{N_\bullet}{V_\bullet},\;\; \alpha = \frac{V_\bullet}{a+b+V_\bullet}.$$

Preostaje nam još odrediti kvadratni gubitak za F^{Bayes}.

$$E\left[(F^{Bayes}-\Theta)^2\right] \;=\; E\left[E\left[(F^{Bayes}-\Theta)^2 \mid \Theta\right]\right] \;=\; \alpha^2 E\left[Var\left[\overline{N} \mid \Theta\right]\right] \;+\; (1\;-\;\alpha)^2\; Var\,[\Theta]\,.$$

Sada ću posebno raspisati izraz za $E\left[Var\left[\overline{N} \mid \Theta\right]\right]$ i onda ću ga vratiti natrag u prethodnu jednadžbu.

$$\begin{aligned} E\left[Var\left[\overline{N} \mid \Theta\right]\right] &= \frac{1}{V_\bullet}E[\Theta(1-\Theta)] = \frac{1}{V_\bullet}\left(E[\Theta] - Var[\Theta] - (E[\Theta])^2\right) \\ &= \frac{1}{V_\bullet}\left(\frac{a}{a+b} - \frac{a^2}{(a+b)^2} - Var[\Theta]\right) = \frac{1}{V_\bullet}\left(\frac{ab}{(a+b)^2} - Var[\Theta]\right) \\ &= \frac{1}{V_\bullet}\left(\frac{ab(1+a+b)}{(a+b)^2(1+a+b)} - \frac{ab}{(1+a+b)(a+b)^2}\right) \\ &= \frac{1}{V_\bullet}(1+a+b-1)\frac{ab}{(1+a+b)(a+b)^2} = \frac{1}{V_\bullet}(a+b)Var[\Theta] \end{aligned}$$

Sada imamo,

$$
\begin{aligned}
E\left[(\widetilde{\Theta}-\Theta)^2\right] &= \alpha^2 \frac{1}{V_\bullet}(a+b)Var[\Theta] \;+\; (1\;-\;\alpha)^2\; Var[\Theta] \\
&= \frac{V_\bullet^2}{(a+b+V_\bullet)^2}\frac{1}{V_\bullet}(a+b)Var[\Theta] \;+\; \left(1-\frac{V_\bullet}{(a+b+V_\bullet)}\right)^2 Var[\Theta] \\
&= \frac{V_\bullet}{(a+b+V_\bullet)^2}(a+b)Var[\Theta] \;+\; \frac{(a+b)^2}{(a+b+V_\bullet)^2}Var[\Theta] \\
&= \left(\frac{V_\bullet}{a+b+V_\bullet} \;+\; \frac{a+b}{a+b+V_\bullet}\right)\frac{a+b}{a+b+V_\bullet}Var[\Theta] = \frac{a+b}{a+b+V_\bullet}Var[\Theta].
\end{aligned}
$$

Dobiveni izraz ćemo koristiti da dokažemo jednakosti 2.26 i 2.27 za kvadratni gubitak F^{Bayes}.

Sada imamo,

$$
\begin{aligned}
E\left[(F^{Bayes}-\Theta)^2\right] &= \frac{a+b}{a+b+V_\bullet}Var[\Theta] = \left(1-\frac{V_\bullet}{a+b+V_\bullet}\right)Var[\Theta] = (1-\alpha)\,Var[\Theta] \\
&= (1-\alpha)\,E\left[(\Theta-E[\Theta])^2\right] = (1-\alpha)\,E\left[(F^{coll}-\Theta)^2\right],
\end{aligned}
$$

time smo dokazali prvu jednakost 2.26.

Pokažimo sada i drugu jednakost 2.27,

$$
\begin{aligned}
E\left[(F^{Bayes}-\Theta)^2\right] &= \frac{a+b}{a+b+V_\bullet}Var[\Theta] = \frac{a+b}{a+b+V_\bullet}\;E\left[Var\left[\overline{N}|\Theta\right]\right]\;V_\bullet\;\frac{1}{a+b} \\
&= \frac{V_\bullet}{a+b+V_\bullet}\;E\left[Var\left[\overline{N}|\Theta\right]\right] = \alpha\;E\left[Var\left[\overline{N}|\Theta\right]\right] \\
&= \alpha\;E\left[E\left[\left(\overline{N}-\Theta\right)^2\;|\;\Theta\right]\right] = \alpha\;E\left[\left(\overline{N}-\Theta\right)^2\right]
\end{aligned}
$$

□

2.3.3 Normalni-normalni model

Promatramo individualni rizik. Neka je $\mathbf{X} = (X_1, X_2, ..., X_n)'$ slučajni vektor, gdje X_j predstavlja iznos štete u godini j za promatranog pojedinca. Želimo odrediti X_{n+1}, tj. njegovu očekivanu štetu u sljedećem periodu.

Pretpostavke normalnog-normalnog modela

- uvjetno, za dano $\Theta = \vartheta$, slučajne varijable X_j, $(j = 1, 2, ..., n)$ su nezavisne i normalno distribuirane, tj.
$$X_j \sim N(\vartheta, \sigma^2)$$
- slučajna varijabla Θ je normalno distribuirana, tj.
$$\Theta \sim N(\mu, \tau^2),$$
strukturna funkcija ima gustoću
$$u(\vartheta) = \frac{1}{\tau\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{\vartheta-\mu}{\tau}\right)^2}$$

Propozicija 2.3.5. *Uz pretpostavku normalnog-normalnog modela, vrijede sljedeće relacije:*

$$P^{ind} = E[N_{n+1} \mid \Theta] = \Theta, \tag{2.28}$$

$$P^{coll} = E[\Theta] = \mu, \tag{2.29}$$

$$P^{Bayes} = \frac{\tau^2\mu + \sigma^2 X_\bullet}{\tau^2 + n\sigma^2} = \alpha\overline{X} + (1 - \alpha)\,\mu, \ \textit{gdje je}\ \ \alpha = \frac{n}{n + \frac{\sigma^2}{\tau^2}}, \ \overline{X} = \frac{1}{n} X_\bullet, \tag{2.30}$$

Kvadratni gubitak za P_{Bayes} dan je sa:

$$\begin{aligned} E[(P^{Bayes} - \Theta)^2] &= (1 - \alpha)E[(P^{coll} - \Theta)^2] \qquad (2.31)\\ &= \alpha E[(\overline{X} - \Theta)^2]. \qquad (2.32) \end{aligned}$$

Dokaz. Kako je uvjetna distribucija od X_{n+1} za dano $\Theta = \vartheta$ normalna $N(\vartheta, \sigma^2)$, vrijedi:

$$P^{ind} = E[X_{n+1}|\Theta] = \Theta.$$

Budući da je $\Theta \sim N(\mu, \tau^2)$ direktno slijedi:

$$P^{coll} = E[\Theta] = \mu.$$

Da bismo odredili $P^{Bayes} = \widetilde{\Theta} = E[\Theta|\mathbf{X}]$, najprije ćemo odrediti *a posteriori* gustoću od slučajne varijable Θ za dano $\mathbf{X} = \mathbf{x}$.

$$u(\vartheta|\mathbf{x}) = \frac{\frac{1}{\tau\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{\vartheta-\mu}{\tau}\right)^2} \prod_{j=1}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_j-\vartheta}{\sigma}\right)^2}}{\int_0^1 \frac{1}{\tau\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{\xi-\mu}{\tau}\right)^2} \prod_{j=1}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_j-\xi}{\sigma}\right)^2} \mathrm{d}\xi},$$

a odavde slijedi

$$u(\vartheta|\mathbf{x}) \;\propto\; \frac{1}{\tau\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{\vartheta-\mu}{\tau}\right)^2} \prod_{j=1}^n \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_j-\vartheta}{\sigma}\right)^2}\right). \tag{2.33}$$

U eksponentu desne strane 2.33, pri čemu ćemo sa β_1 i β_2 označiti one dijelove koji ne ovise o ϑ, imamo

$$\begin{aligned}
&-\frac{1}{2}\left(\left(\tau^{-2} + n\sigma^{-2}\right)\vartheta^2 - 2\left(\tau^{-2}\mu + n\sigma^{-2}\overline{x}\right)\vartheta + \beta_1\right) \\
&= -\frac{1}{2}\left(\tau^{-2} + n\sigma^{-2}\right)\left(\left(\vartheta - \frac{\tau^{-2}\mu + n\sigma^{-2}\overline{x}}{\tau^{-2} + n\sigma^{-2}}\right)^2 + \beta_2\right) \\
&= -\frac{1}{2}\left(\tau^{-2} + n\sigma^{-2}\right)\left(\left(\vartheta - \frac{\sigma^2\mu + n\tau^2\overline{x}}{\sigma^2 + n\tau^2}\right)^2 + \beta_2\right)
\end{aligned}$$

Dakle, *a posteriori* distribucija slučajne varijable Θ, za dano $\mathbf{X}$, ponovno je normalna distribucija, ali sada sa novim parametrima

$$\mu' = \frac{\sigma^2\mu + n\tau^2\overline{X}}{\sigma^2 + n\tau^2} \quad \text{i} \quad (\tau')^2 = \frac{\tau^2\sigma^2}{\sigma^2 + n\tau^2}.$$

Time smo pokazali da je

$$P^{Bayes} = \widetilde{\Theta} = E[\Theta|\mathbf{X}] = \mu' = \frac{\sigma^2\mu + n\tau^2\overline{X}}{\sigma^2 + n\tau^2}.$$

Sada ćemo dokazati i drugu jednakost u izrazu za P^{Bayes}.

$$\begin{aligned}
\mu' &= \frac{\sigma^2\mu + n\tau^2\overline{X}}{\sigma^2 + n\tau^2} = \frac{n\tau^2}{\sigma^2 + n\tau^2}\overline{X} + \frac{\sigma^2\overline{X}}{\sigma^2 + n\tau^2}\mu = \frac{n}{n + \frac{\sigma^2}{\tau^2}}\overline{X} + \left(\frac{\sigma^2 + n\tau^2}{\sigma^2 + n\tau^2} - \frac{n\tau^2}{\sigma^2 + n\tau^2}\right)\mu \\
&= \frac{n}{n + \frac{\sigma^2}{\tau^2}}\overline{X} + \left(1 - \frac{n}{n + \frac{\sigma^2}{\tau^2}}\right)\mu = \alpha\overline{X} + (1 - \alpha)\mu, \;\; \text{gdje je } \alpha = \frac{n}{n + \frac{\sigma^2}{\tau^2}}.
\end{aligned}$$

Preostaje nam još odrediti kvadratni gubitak za P^{Bayes}.

$$\begin{aligned} E\left[(\widetilde{\Theta}-\Theta)^2\right] &= E\left[E\left[(\widetilde{\Theta}-\Theta)^2 \mid \Theta\right]\right] = E\left[Var\left[\Theta|\mathbf{X}\right]\right] = (\tau')^2 = \tau^2\frac{\sigma^2}{\sigma^2+n\tau^2} \\ &= \tau^2\left(1-\frac{n\tau^2}{\sigma^2+n\tau^2}\right) = \tau^2\left(1-\frac{n}{n+\frac{\sigma^2}{\tau^2}}\right) = \tau^2\left(1-\alpha\right) = Var[\Theta]\left(1-\alpha\right) \\ &= E\left[(\Theta-E[\Theta])^2\right](1-\alpha) = (1-\alpha)\,E\left[\left(P^{coll}-\Theta\right)^2\right]. \end{aligned}$$

Time smo dokazali jednakost 2.31.

Pokažimo sada i jednakost 2.32.

$$E\left[(\widetilde{\Theta}-\Theta)^2\right] = (1-\alpha)\tau^2 = \left(1-\frac{n}{n+\frac{\sigma^2}{\tau^2}}\right)\tau^2 = \frac{\frac{\sigma^2}{\tau^2}}{n+\frac{\sigma^2}{\tau^2}}\,\tau^2 = \frac{n}{n+\frac{\sigma^2}{\tau^2}}\frac{1}{n}\,\sigma^2 = \alpha\,\frac{\sigma^2}{n}.$$

Budući da vrijedi

$$\begin{aligned} E\left[(\overline{X}-\Theta)^2\right] &= E\left[E\left[\left(\overline{X}-\Theta\right)^2|\Theta\right]\right] = E\left[Var\left[\overline{X}|\Theta\right]\right] = \frac{1}{n^2}E\left[Var\left[\sum_{j=1}^{n}X_j|\Theta\right]\right] \\ &= \frac{1}{n^2}E\left[\sum_{j=1}^{n}Var\left[X_j|\Theta\right]\right] = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}, \end{aligned}$$

imamo

$$E\left[(\widetilde{\Theta}-\Theta)^2\right] = \alpha\,E\left[(\overline{X}-\Theta)^2\right]$$

što smo i htjeli dokazati. □

Primjer 2.3.6 (Numerička ilustracija normalnog-normalnog modela)**.** *Neka je* $\mathbf{X} = (X_1, X_2, ..., X_n)'$ *slučajni vektor, gdje* X_j *predstavlja iznos štete u godini* j *za promatranog pojedinca. Želimo odrediti* X_{n+1}, *tj. njegovu očekivanu štetu u sljedećem periodu.*

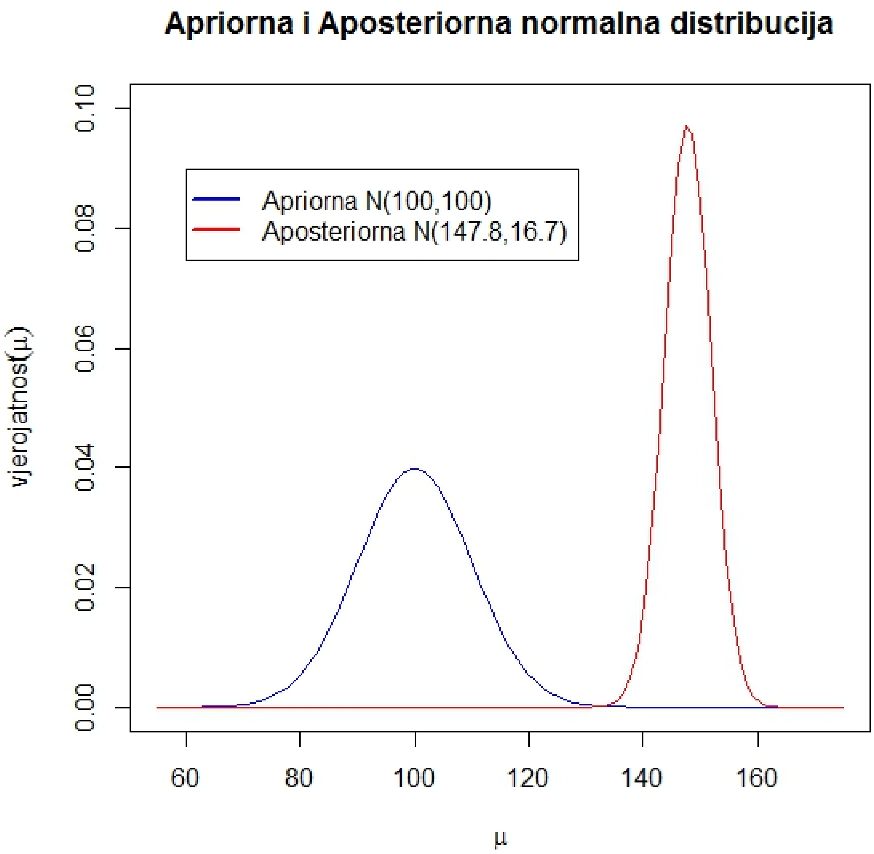
Pretpostavimo da u 10 sukcesivnih godina iznosi šteta imaju sljedeće vrijednosti:

$$3,\ 46.3,\ 70.9,\ 79.8,\ 92.4,\ 206.1,\ 225.2,\ 234.1,\ 275.7,\ 340.$$

Ovi brojevi dolaze iz normalnog-normalnog modela uz (brojevi su simulirani pa znamo da dolaze iz $N(\vartheta, 200)$, $\vartheta = 100$*) pretpostavku da je apriori razdioba za parametar* ϑ $N(100, 100)$.

Tada prema formuli a posteriori razdioba od ϑ *je* $N(147.8, 278.89)$

Grafički prikaz:



Bayesovski procjenitelj za premiju dan formulom

$$E[\Theta|X_1, X_2, \ldots, X_n] = \frac{\tau^2\mu \ + \ \sigma^2 X_\bullet}{\tau^2 \ + \ n\sigma^2}.$$

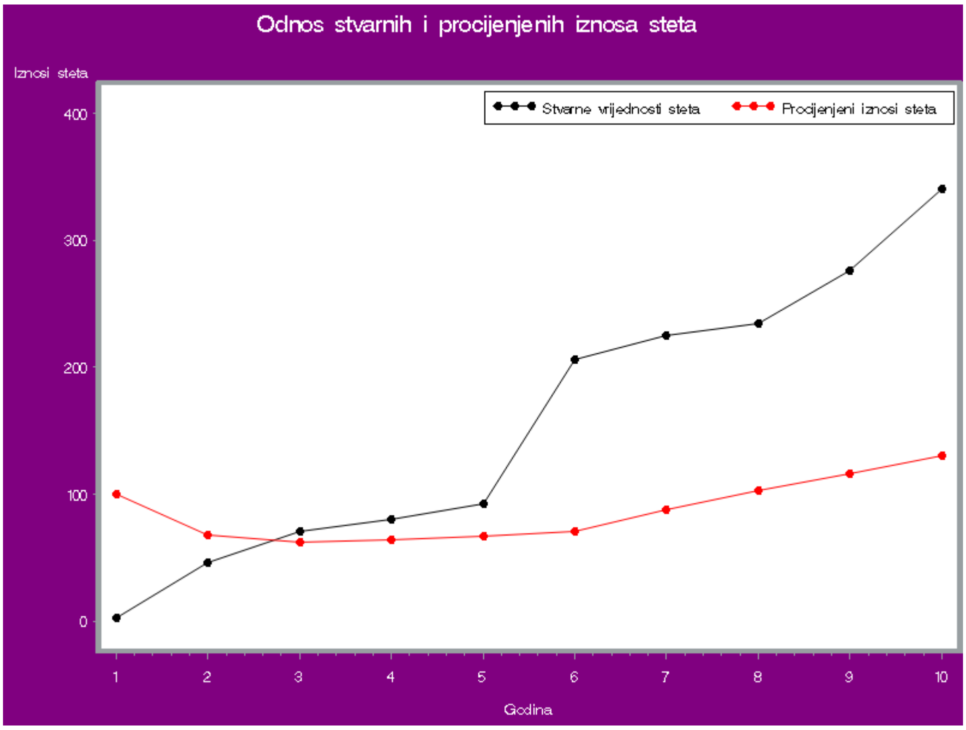
Procjena očekivanog ukupnog iznosa šteta u sljedećoj (11-stoj) godini iznosi

$$E[\Theta|X_1, X_2, \ldots, X_n] = 147.9$$

Rezulate procijenjenih iznosa šteta u ovisnosti distribucije za parametar ϑ *koje sam dobila računom prikazat ću u sljedećoj tablici:*

godina	stvarne vrijednosti šteta	procijenjeni iznosi šteta za $\mathbb{N}(100, 100)$
1	3	100
2	46.3	67.66
3	70.9	62.3
4	79.8	64.04
5	92.4	66.7
6	206.1	70.3
7	225.2	87.3
8	234.1	102.6
9	275.7	115.78
10	340	130.3

Grafički prikaz rezultata:



Slika 2.5: Odnosi između stvarnih vrijednosti šteta i procijenjenih po godinama

Slika 2.5 pokazuje da Bayesovski procjenitelj sporo konvergira k stvarnoj vrijednosti za naše pretpostavke.

2.3.4 Zajedničke karakteristike triju promatranih slučaja

U sva tri slučaja,

- Bayesova premija je konveksna funkcija promatranih veličina, a time i premija povjerenja.
- P^{Bayes} možemo izraziti u obliku:

$$P^{Bayes} = \alpha\,\overline{X} + (1 - \alpha)\,P^{coll}.$$

- težina α je dana sa:

$$\alpha = \frac{n}{n + \kappa}, \quad \text{gdje je } \kappa \text{ pripadna konstanta.}$$

- kvadratni gubitak za Bayesovu premiju dan je sa:

$$E[(P^{Bayes} - \Theta)^2] = (1 - \alpha)E[(P^{coll} - \Theta)^2] = \alpha E[(\overline{X} - \Theta)^2].$$

- pokazali smo da *a posteriori* distribucija od slučajne varijable Θ pripada istoj familiji kao i *a priori* distribucija.

Poglavlje 3

Procjenitelji povjerenja

Vidjeli smo da je Bayesova premija

$$\widetilde{\mu(\Theta)} = E[\mu(\Theta) \mid \mathbf{X}]$$

najbolji mogući procjenitelj u klasi svih procjenitelja. Bayesov procjenitelj često ne možemo izraziti u zatvorenom analitičkom obliku i možemo ga računati samo numeričkim metodama. Štoviše, da bismo izračunali $\widetilde{\mu(\Theta)}$, moramo odrediti uvjetne distribucije kao i *a priori* distribucije šteta, što se u praksi često ne može zaključiti niti iz danih podataka niti pogoditi intuicijom. Stoga ćemo pokušati umjesto Bayesovog procjenitelja naći procjenitelj za $\mu(\Theta)$ među linearnim funkcijama našeg uzorka $\mathbf{X} = (X_1, X_2, ..., X_n)'$, tj. tražit ćemo najbolje procjenitelje u srednjekvadratnom smislu u klasi svih linearnih procjenitelja.

3.1 Procjenitelji povjerenja u jednostavnom kontekstu

3.1.1 Premija povjerenja u jednostavnom modelu povjerenja

Pretpostavke jednostavnog modela povjerenja

- slučajne varijable X_j, $j = 1, 2, ..., n$, uz uvjet $\Theta = \vartheta$, su nezavisne i jednako distribuirane s funkcijom distribucije F_ϑ i sa uvjetnim momemtima

$$\mu(\vartheta) = E[X_j \mid \Theta = \vartheta], \ \ \sigma^2(\vartheta) = Var[X_j \mid \Theta = \vartheta].$$

- Θ je slučajna varijabla sa funkcijom ditribucije $U(\vartheta)$.

Par (Θ, X_j), $j = 1, 2, ..., n$, opisuje pojedinu policu u portfelju, gdje je Θ nepoznat, a niz slučajnih varijabli (X_j) predstavlja niz šteta u danoj polici. Budući da su za dano $\Theta = \vartheta$ slučajne varijable X_j, $j = 1, 2, ..., n$ jednako distribuirane, momenti $\mu(\vartheta)$ i $\sigma^2(\vartheta)$ ne ovise o indeksu j. Nadalje, primijetimo da su, s obzirom na to da je Θ slučajna varijabla, $\mu(\Theta)$ i $\sigma^2(\Theta)$ također slučajne varijable.

U ovom modelu imamo

$$\begin{aligned} P^{ind} &= \mu(\Theta) = E[X_{n+1} \mid \Theta], \\ P^{coll} &= \mu_0 = \int_\Theta \mu(\vartheta)\, dU(\vartheta). \end{aligned}$$

Ponovno nam je cilj procijeniti individualnu premiju $\mu(\Theta)$ uz dano $\mathbf{X}$, no sada ćemo procjenitelj za $\mu(\Theta)$ pokušati naći među linearnim funkcijama našeg uzorka. Najbolji procjenitelj unutar klase svih linearnih procjenitelja ćemo označiti sa P^{cred} ili $\widehat{\widehat{\mu(\Theta)}}$ i zvat ćemo ga **procjenitelj povjerenja**.

Prema definiciji, $\widehat{\widehat{\mu(\Theta)}}$ ima oblik

$$\widehat{\widehat{\mu(\Theta)}} = \widehat{a_0} + \sum_{j=1}^{n} \widehat{a_j} X_j,$$

gdje su $\widehat{a_0}, \widehat{a_1}, ..., \widehat{a_n}$ rješenja problema minimizacije

$$E\left[\left(\mu(\Theta) - \widehat{a_0} - \sum_{j=1}^{n} \widehat{a_j} X_j\right)^2\right] = \min_{a_0, a_1, \ldots, a_n \in \mathbb{R}} E\left[\left(\mu(\Theta) - a_0 - \sum_{j=1}^{n} a_j X_j\right)^2\right].$$

Parcijalna derivacija po a_0 nam daje sljedeću jednadžbu

$$E\left[\mu(\Theta) - a_0 - \sum_{j=1}^{n} a_j X_j\right] = 0$$

pa iz

$$E[X_j] = E[E[X_j|\Theta]] = E[\mu(\Theta)]$$

dobijemo

$$a_0 = E[\mu(\Theta)]\left(1 - \sum_{j=1}^{n} a_j\right). \tag{3.1}$$

Ako diferenciramo srednjekvadratnu grešku po $a_i,\ i > 0$, dobijemo

$$E\left[X_i\left(\mu(\Theta) - a_0 - \sum_{j=1}^{n} a_j\right)\right] = 0. \tag{3.2}$$

Uočimo da vrijedi

$$E[X_i\mu(\Theta)] = E[E[X_i\mu(\Theta)|\Theta]] = E[\mu(\Theta)^2] = Var[\mu(\Theta)] + (E[\mu(\Theta)])^2 = \tau^2 + \mu_0^2$$

gdje su

$$\tau := Var[\mu(\Theta)],$$
$$\mu_0 := E[\mu(\Theta)].$$

Slično vrijedi i za $j \neq k$

$$E[X_jX_k] = Var[\mu(\Theta)] + (E[\mu(\Theta)])^2 = \tau^2 + \mu_0^2.$$

Na kraju uočimo da vrijedi

$$\begin{aligned} E[X_i^2] &= E[E[X_i^2|\Theta]] = E[Var[X_i|\Theta]] + E[\mu(\Theta)^2] \\ &= E[Var[X_i|\Theta]] + Var[\mu(\Theta)] + (E[\mu(\Theta)^2])^2 \\ &= \sigma^2 + \tau^2 + \mu_0^2, \end{aligned}$$

gdje je

$$\sigma^2 := E[Var[X_i|\Theta]].$$

Sada 3.2 možemo zapisati kao

$$a_i\sigma^2 = \left(1 - \sum_{i=1}^{n} a_i\right)\left(\tau^2 + \mu_0^2 - a_0\mu_0\right). \tag{3.3}$$

Odavde je jasno da mora vrijediti da je $a_1 = a_2 = ... = a_n$.

Označimo sa $\widehat{b} = \sum_{i=1}^{n} a_i$, tj. $\widehat{a_i} = \frac{\widehat{b}}{n}$ za $i > 0$ i sa $\widehat{a} = a_0$. Sada traženi procjenitelj možemo zapisati u obliku

$$\widehat{\widehat{\mu(\Theta)}} = \widehat{a} + \widehat{b}\overline{X},$$

gdje je

$$\overline{X} = \frac{1}{n}\sum_{j=1}^{n} X_j.$$

Iz jednadžbi 3.1 i 3.3 dobijemo

$$\widehat{a} = (1 - \widehat{b})\mu_0$$
$$\widehat{b} = \frac{\tau^2}{\tau^2 + \frac{\sigma^2}{n}} = \frac{n}{n + \frac{\sigma^2}{\tau^2}}.$$

Time smo dokazali sljedeći teorem:

Teorem 3.1.1. *Uz pretpostavke jednostavnog modela povjerenja, procjenitelj povjerenja dan je sa*

$$\widehat{\widehat{\mu(\Theta)}} = \alpha\overline{X} + (1-\alpha)\mu_0, \tag{3.4}$$

gdje su

$$\mu_0 = E[\mu(\Theta)], \tag{3.5}$$

$$\alpha = \frac{n}{n + \frac{\sigma^2}{\tau^2}}. \tag{3.6}$$

Napomena

- Parametre σ^2, τ^2 i μ_0 tretirali smo kao unaprijed poznate iako to u praksi nije slučaj. U bayesovskim primjerima, kao npr. u normalnom-normalnom modelu, ti parametri zaista jesu poznati jer nam je poznata *a priori* razdioba. Zovemo ih **strukturalni parametri**.
- Kvocijent $\kappa = \frac{\sigma^2}{\tau^2}$ zovemo **koeficijent povjerenja**. Možemo ga pisati i kao $\kappa = (\frac{\sigma}{\mu_0})^2(\frac{\tau}{\mu_0})^{-2}$. Koeficijent $\frac{\tau}{\mu_0}$ je dobra mjera heterogenosti portfelja, a $\frac{\sigma}{\mu_0} = \frac{\sqrt{E[Var[X_j|\Theta]]}}{E[X_j]}$ je očekivana standardna devijacija unutar rizika podijeljena ukupnim očekivanim iznosom šteta za promatranog pojedinaca u godini j, što je dobra mjera varijabilnosti unutar rizika.
- **težina povjerenja** α povećava se kako se:
  - broj promatranih godina n povećava,
  - koeficijent $\frac{\tau}{\mu_0}$ povećava,
  - koeficijent $\frac{\sigma}{\mu_0}$ smanjuje.

3.1.2 Kvadratni gubitak premije povjerenja

Premiju povjerenja P^{cred}, kao težinski prosjek od P^{coll} i $\overline{X}$, možemo interpretirati i na sljedeći način:

- $P^{coll} = \mu_0$ je najbolji procjenitelj temeljen samo na *a priori* znanju. Njegov kvadratni gubitak je
$$E\left[(\mu_0 - \mu(\Theta))^2\right] = Var\left(\mu(\Theta)\right) = \tau^2.$$
- $\overline{X}$ je najbolji mogući linearni i nepristran procjenitelj (tj. uvjetno nepristan za dano $\Theta = \vartheta$) temeljen samo na vektoru $\mathbf{X}$. Njegov kvadratni gubitak je
$$E\left[\left(\overline{X} - \mu(\Theta)\right)^2\right] = E\left[\frac{\sigma^2(\Theta)}{n}\right] = \frac{\sigma^2}{n}.$$

- $P^{cred} = \widehat{\widehat{\mu(\Theta)}}$ je težinski prosjek ta dva procjenitelja gdje su težine proporcionalne inverzu kvadratnog gubitka

$$\widehat{\widehat{\mu(\Theta)}} = \alpha\overline{X} + (1-\alpha)\mu_0,$$

gdje je

$$\alpha = \frac{\frac{n}{\sigma^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}} = \frac{\tau^2}{\tau^2+\frac{\sigma^2}{n}} = \frac{n}{n+\frac{\sigma^2}{\tau^2}},$$

a kvadratni gubitak ovog procjenitelja je dan u sljedećem teoremu:

Teorem 3.1.2. *Kvadratni gubitak procjenitelja povjerenja $\widehat{\widehat{\mu(\Theta)}}$ dan je s*

$$E\left[\left(\widehat{\widehat{\mu(\Theta)}} - \mu(\Theta)\right)^2\right] = (1-\alpha)\,\tau^2 \tag{3.7}$$

$$= \alpha\frac{\sigma^2}{n} \tag{3.8}$$

Dokaz.

$$\begin{aligned}
E\left[\left(\widehat{\widehat{\mu(\Theta)}} - \mu(\Theta)\right)^2\right] &= E\left[\left(\alpha\left(\mu(\Theta) - \overline{X}\right) + (1-\alpha)\left(\mu(\Theta)-\mu_0\right)\right)^2\right] \\
&= \alpha^2\frac{\sigma^2}{n} + (1-\alpha)^2\,\tau^2 = \left(\frac{n}{n+\frac{\sigma^2}{\tau^2}}\right)^2\frac{\sigma^2}{n} + \left(\frac{\frac{\sigma^2}{\tau^2}}{n+\frac{\sigma^2}{\tau^2}}\right)^2\tau^2 \\
&= \frac{\sigma^2}{n+\frac{\sigma^2}{\tau^2}} = \alpha\frac{\sigma^2}{n} \quad \text{(dokazali smo jednakost 3.7)} \\
&= \frac{\frac{\sigma^2}{\tau^2}}{n+\frac{\sigma^2}{\tau^2}}\,\tau^2 = (1-\alpha)\,\tau^2 \quad \text{(dokazali smo jednakost 3.8)}\ .
\end{aligned}$$

□

Sljedeći korolar daje nam pregled kvadratnih gubitaka za različite procjenitelje za $\mu(\Theta)$.

Korolar 3.1.3. *Uz pretpostavke jednostavnog modela povjerenja imamo:*

	PROCJENITELJ	KVADRATNI GUBITAK
a	$P^{coll} = \mu_0$	$Var(\mu(\Theta))$
b	$P^{cred} = \widehat{\widehat{\mu(\Theta)}}$	$(1-\alpha)Var(\mu(\Theta))$
c	$P^{Bayes} = \widetilde{\mu(\Theta)}$	$E[Var(\mu(\Theta) \mid \mathbf{X})]$

Za kvadratne gubitke vrijedi sljedeći uređaj: a) $\geq$ b) $\geq$ c). Uočimo da je kvadratni gubitak procjenitelja povjerenja veći od kvadratnog gubitka Bayesovog procjenitelja, no to nam je prihvatljivo jer je premiju povjerenja jednostavnije izračunati i za nju nam nije potrebna cijela *a priori* distribucija.

3.1.3 Jednostavan Bülhmannov model i homogeni procjenitelj povjerenja

Do sada smo promatrali samo jedan određeni rizik i tražili smo procjenitelj povjerenja koji se temelji samo na podacima za taj određeni rizik. U praksi obično imamo cijeli portfelj sličnih rizika $i = 1, 2, ..., I$, odnosno sličnih polica osiguranja. Stoga ćemo pretpostaviti da imamo podatke o više polica.

i-tu policu opisuje par $(\Theta_i, \mathbf{X}_i)$, gdje je $\mathbf{X}_i = (X_{i1}, X_{i2}, ..., X_{in})'$ niz šteta u danoj polici, tj. niz šteta rizika i, a Θ_i njegov profil rizika. Za svaki rizik i portfelja, pretpostavit ćemo da Θ_i i $\mathbf{X}_i$ zadovoljavaju pretpostavke jednostavnog modela povjerenja.

Pretpostavke jednostavnog Bülhmannovog modela

B1: Slučajne varijable $X_{ij},\ j = 1, 2, ..., n$ su, uz uvjet $\Theta_i = \vartheta$, nezavisne i jednako distribuirane s funkcijom distribucije F_ϑ i sa uvjetnim momentima

$$\mu(\vartheta) = E[X_{ij} \mid \Theta_i = \vartheta],\ \ \sigma^2(\vartheta) = Var[X_{ij} \mid \Theta_i = \vartheta].$$

B2: Parovi $(\Theta_1, \mathbf{X}_1), (\Theta_2, \mathbf{X}_2), ..., (\Theta_I, \mathbf{X}_I)$ su nezavisni i jednako distribuirani.

Napomena

- Ovaj model se razlikuje od dosad promatranog modela u tome što nam više nije poznata distribucija slučajnih varijabli Θ_i. Kako su za dano $\Theta_i = \vartheta$ slučajne varijable $X_{ij},\ j = 1, 2, ..., n$ jednako distribuirane, $\mu(\vartheta)$ i $\sigma^2(\vartheta)$ ne ovise o indeksu j.
- Pod pretpostavkom B2, modeliran je heterogen portfelj. Profili rizika $\Theta_1, \Theta_2, ..., \Theta_I$ su nezavisne slučajne varijable izabrane iz iste urne sa zakonom razdiobe $U(\vartheta)$. Dakle rizik portfelja ima različite profile rizika. Međutim, rizici portfelja imaju nešto i zajedničko, *a priori* su jednaki, tj. *a priori* ne mogu biti prepoznati kao različiti.

Cilj nam je pronaći procjenitelj povjerenja u jednostavnom Bülhmannovom modelu. Budući da za svaki rizik i portfelja, želimo procijeniti njegovu individualnu premiju $\mu(\Theta_i)$, imat ćemo procjenitelje povjerenja $\widehat{\widehat{\mu(\Theta_i)}}$ od $\mu(\Theta_i)$ za $i = 1, 2, ..., I$.

Procjenitelj povjerenja definiramo kao najbolji procjenitelj koji je linearna funkcija svih podataka u portfelju, tj. procjenitelj povjerenja $\widehat{\widehat{\mu(\Theta_i)}}$ od $\mu(\Theta_i)$ je po definiciji najbolji procjenitelj u klasi

$$\left\{ a \; + \; \sum_{k=1}^{I} \sum_{j=1}^{n} b_{kj} \, X_{kj} \; : \; a, b_{kj} \in \mathbb{R} \right\}.$$

Za dani portfelj rizika, možemo promatrati još jedan tip procjenitelja povjerenja od $\mu(\Theta_i)$ kojeg zovemo **homogeni procjenitelj povjerenja**. Homogeni procjenitelj povjerenja, u oznaci $\widehat{\widehat{\mu(\Theta_i)}}^{hom}$, definiramo kao najboljeg procjenitelja od $\mu(\Theta_i)$ u klasi nepristranih procjenitelja

$$\left\{ \widehat{\mu(\Theta_i)} \; = \sum_{k=1}^{I} \sum_{j=1}^{n} b_{kj} \, X_{kj} \; : \; E\left[\widehat{\mu(\Theta_i)}\right] = E[\mu(\Theta_i)], b_{kj} \in \mathbb{R} \right\}.$$

Uočimo da za razliku od procjenitelja povjerenja $\widehat{\mu(\Theta_i)}$, ne sadrži konstantni član i da se zahtjeva da bude nepristran procjenitelj, što je inače automatski ispunjen uvjet za nehomogeni procjenitelj povjerenja.

Neka su

$$\begin{aligned} \tau^2 &:= Var(\mu(\Theta)), \\ \sigma^2 &:= E[Var[X|\Theta]]. \end{aligned}$$

Teorem 3.1.4. *(Nehomogeni) procjenitelja povjerenja $\widehat{\widehat{\mu(\Theta_i)}}$ i homogenn procjenitelj povjerenja $\widehat{\widehat{\mu(\Theta_i)}}^{hom}$ u jednostavnom Bülhmannovom modelu dani su s*

$$\widehat{\widehat{\mu(\Theta_i)}} = \alpha \overline{X_i} + (1 - \alpha) \, \mu_0 \tag{3.9}$$

$$\widehat{\widehat{\mu(\Theta_i)}}^{hom} = \alpha \overline{X_i} + (1 - \alpha) \, \overline{X} \tag{3.10}$$

gdje su

$$\begin{aligned} \alpha &= \frac{n}{n + \frac{\sigma^2}{\tau^2}}, \quad \mu_0 = E[\mu(\Theta)], \\ \overline{X_i} &= \frac{1}{n} \sum_{j=1}^{n} X_{ij}, \quad \overline{X} = \frac{1}{In} \sum_{i=1}^{I} \sum_{j=1}^{n} X_{ij}. \end{aligned}$$

Napomena
Uočimo da je α rastuća funkcija od n, tj. duljine uzorka. S druge strane, α pada kako raste koeficijent povjerenja $\frac{\sigma^2}{\tau^2}$. Koeficijent povjerenja mjeri odnos varijabilnosti procjenitelja za $\mu(\Theta_i)$, odnosno za konkretnu policu i u odnosu na varijabilnost procjenitelja za sve rizike u kolektivu.

Dokaz teorema 3.1.4. Dokaz jednadžbe 3.9 analogan je dokazu teorema 3.1.1. Da bi dokazali jednadžbu 3.10 za homogeni procjenitelj povjerenja $\widehat{\widehat{\mu(\Theta_i)}}^{hom}$, minimizirat ćemo sljedeći izraz

$$E\left[\left(\mu(\Theta_i) - \sum_{k=1}^{I}\sum_{j=1}^{n} b_{kj}X_{kj}\right)^2\right]$$

uvjetno na zahtjev nepristranosti da je

$$E\left[\sum_{k=1}^{I}\sum_{j=1}^{n} b_{kj}X_{kj}\right] = \mu_0.$$

Iz zahtjeva nepristranosti i činjenice da je

$$E[\mu(\Theta_i)] = E[X_{kj}] = \mu_0, \forall\ i, j, k\ ,$$

dobijemo uvjet

$$\sum_{k=1}^{I}\sum_{j=1}^{n} b_{kj} = 1.$$

Umjesto formalnog rješavanja problema minimizacije, malo ćemo bolje pogledati $\widehat{\widehat{\mu(\Theta_i)}}$ koji je definiran jednadžbom 3.9. Da bi taj (nehomogeni) procjenitelj povjerenja transformirali u homogeni, konstantni član $(1 - \alpha)\mu_0$ moramo zamijeniti prikladnom linearnom funkcijom svih podataka. Prirodno je μ_0 zamijeniti uzoračkom aritmetičkom sredinom

$$\overline{X} = \frac{1}{In}\sum_{i=1}^{I}\sum_{j=1}^{n} X_{ij}.$$

Lako se provjeri da je taj procjenitelj zaista rješenje problema minimizacije, a zbog toga i homogeni procjenitelj povjerenja. $\square$

Poglavlje 4

Bülhmann-Straubov model

Bülhmannov model dalje su profinili Bülhmann i Straub. Njihova osnovna ideja je bila dopustiti heterogenost unutar svake pojedinačne police osiguranja. To je najvažniji model povjerenja u osiguranju.

4.1 Pretpostavke modela

Dan nam je portfelj sa I rizika ili "kategorija rizika", odnosno dan nam je portfelj sa I polica. Koristimo sljedeće veličine za koje ćemo pretpostavit da su nam poznate:

S_{ij} - ukupan iznos štete u godini j,

V_{ij} - volumen rizika i u godini j (npr. broj prodanih polica za rizik i u godini j).

Napomena

- Iz prethodno definiranih veličina dobijemo novi niz slučajnih varijabli

$$X_{ij} = \frac{S_{ij}}{V_{ij}}, \quad j = 1, 2, \ldots ,$$

  gdje X_{ij} predstavlja prosječan iznos štete rizika i u godini j, (eng. *claims ratio*). X_{ij} imaju sličnije razdiobe nego S_{ij}.

- Umjesto oznake V_{ij}, koristit ćemo oznaku w_{ij} koja će nam označavati ukupne (poznate) težine. Težine w_{ij} ćemo općenito interpretirati kao volumen rizika i u godini j, no moguće su i druge interpretacije pa zbog toga koristimo općenitiju oznaku w_{ij} umjesto V_{ij}.

Pretpostavke Bülhmann-Straubovog modela

Rizik i je karakteriziran individualnim profilom rizika ϑ_i, što je realizacija slučajne varijable Θ_i i vrijedi:

BS1: Uvjetno, za dano $\Theta_i = \vartheta$, $\{ X_{ij} : \ j = 1, 2, ..., n\}$ su nezavisne sa

$$E[X_{ij} \mid \Theta_i = \vartheta] = \mu(\vartheta), \tag{4.1}$$

$$Var[X_{ij} \mid \Theta_i = \vartheta] = \frac{\sigma^2(\vartheta)}{w_{ij}}, \ \text{ gdje je } \ \sigma^2(\vartheta) := Var[X_{ij} \mid \Theta_i = \vartheta]. \tag{4.2}$$

BS2: Parovi $(\Theta_1\ ,\ \mathbf{X}_1), (\Theta_2\ ,\ \mathbf{X}_2), ...$ su nezavisni, i $\Theta_1, \Theta_2, ...$ su nezavisne i jednako distribuirane slučajne varijable.

Napomena

- Par $(\Theta_i\ ,\ \mathbf{X}_i)$ opisuje pojedinu policu u portfelju, gdje je Θ_i nepoznat, a niz slučajnih varijabli $\mathbf{X}_i = (X_{i1}, X_{i2}, ..., X_{in})'$ predstavlja niz šteta u danoj polici.

- Primijetimo da su rizici nezavisni, tj. slučajne varijable koje pripadaju različitim policama su nezavisne i ta nam pretpostavka obično ne predstavlja problem u praksi. No pretpostavka da su rizici jednako distribuirani u praksi ne vrijedi uvijek pa se u konkretnim situacijama pitamo da li uopće ima smisla. Na primjer, u osiguranju od požara, odmah nam je jasno da drvodjelačka radnja predstavlja veći rizik od poslovne zgrade.

- Bülhmann-Straubov model se može interpretirati kao model dvije urne (slika 1.1 na stranici 5). Iz prve urne odabiremo profile rizika Θ_i, koji nam određuju sadržaj druge urne. U drugom koraku odabiremo slučajne varijable X_{ij}. Na taj način smo modelirali heterogeni portfelj. Dakle, rizici u portfelju imaju različite profile rizika, ali imaju i zajedničko to da smo profile rizika Θ_i odabirali iz iste urne.

4.2 Premija povjerenja

Za svaki rizik i želimo procijeniti njegovu individualnu premiju rizika $\mu(\Theta_i)$.
Dani su nam podaci $D = \{\mathbf{X}_i : \ i = 1, 2, ..., I\}$, gdje je $\mathbf{X}_i = (X_{i1}, X_{i2}, ..., X_{in})'$ vektor i-tog rizika. Tražimo procjenitelj povjerenja koji se temelji na podacima D.
Sljedeće veličine od interesa su:

RIZIK i	INTERPRETACIJA
$\mu(\Theta_i) = E[X_{ij}\vert\Theta_i]$	individualna premija rizika
$\sigma^2(\Theta_i) = w_{ij}Var[X_{ij}\vert\Theta_i]$	varijanca podataka individualnog rizika

KOLEKTIV/PORTFELJ	INTERPRETACIJA
$\mu_0 = E[\mu(\Theta_i)]$	kolektivna premija
$\sigma^2 = E[\sigma^2(\Theta_i)]$	prosjek varijance individualnog rizika
$\tau^2 = Var[\mu(\Theta_i)]$	varijanca između individualnih premija rizika

Teorem 4.2.1 (procjenitelj povjerenja)**.** *(Nehomogeni) procjenitelj povjerenja u Bülhmann-Straubovom modelu (uz pretpostavke Bülhmann-Straubovog modela) je dan s*

$$\widehat{\widehat{\mu(\Theta_i)}} = \alpha_i\overline{X}_{i\bullet} + (1-\alpha_i)\mu_0 = \mu_0 + \alpha_i(\overline{X}_{i\bullet} - \mu_0), \tag{4.3}$$

gdje su

$$\overline{X}_{i\bullet} := \sum_j \frac{w_{ij}}{w_{i\bullet}} X_{ij}, \quad w_{i\bullet} := \sum_j w_{ij},$$
$$\alpha_i := \frac{w_{i\bullet}}{w_{i\bullet} + \frac{\sigma^2}{\tau^2}} = \frac{w_{i\bullet}}{w_{i\bullet} + \kappa}.$$

Dokaz. Slično kao i u dokazu teorema 3.1.1 tražimo linearni procjenitelj za $\mu(\Theta_i)$ u srednjekvadratnom smislu. Drugim riječima, tražimo konstante $a_0, a_1, ..., a_n$ tako da minimiziramo sljedeći izraz

$$E\left[\left(\mu(\Theta) - a_0 - \sum_{j=1}^{n} a_j X_j\right)^2\right].$$

Na isti način odredimo konstante i na kraju dobijemo

$$a_0 = \frac{\mu_0\frac{\sigma^2}{\tau^2}}{w_{i\bullet} + \frac{\sigma^2}{\tau^2}} \quad \text{i} \quad a_i = \frac{w_{ij}}{w_{ij} + \frac{\sigma^2}{\tau^2}}, \quad i = 1, 2, \ldots .$$

Sada traženi procjenitelj povjerenja možemo zapisati kao

$$\alpha_i\overline{X}_{i\bullet} + (1 - \alpha_i)\mu_0,$$

uz korištenje težinskog prosjeka

$$\overline{X}_{i\bullet} = \sum_j \frac{w_{ij}}{w_{i\bullet}} X_{ij} \quad \text{i} \quad \alpha_i = \frac{w_{i\bullet}}{w_{i\bullet} + \frac{\sigma^2}{\tau^2}} = \frac{w_{i\bullet}}{w_{i\bullet} + \kappa}.$$

□

Napomena

- Uočimo da koeficijenti a_i, $i = 1, 2, ...$, nisu nužno jednaki, osim u slučaju kada su volumeni rizika jednaki.
- $\kappa = \frac{\sigma^2}{\tau^2}$ je koeficijent povjerenja.

4.3 Interpretacija i kvadratni gubitak procjenitelja povjerenja

Interpretacija procjenitelja povjerenja u teoremu 4.2.1:

- μ_0 je najbolji procjenitelj koji se temelji samo na *a priori* znanju i njegov kvadratni gubitak je
$$E\left[(\mu_0 - \mu(\Theta_i))^2\right] = Var[(\mu(\Theta_i))] = \tau^2.$$
- $\overline{X}_{i\bullet}$ je najbolji linearan i nepristran procjenitelj koji se temelji samo na vektoru podataka $\mathbf{X}_i$ i njegov kvadratni gubitak je
$$\begin{aligned} E\left[(\overline{X}_{i\bullet} - \mu(\Theta_i))^2\right] &= E\left\{E\left[\left(\sum_j \frac{w_{ij}}{w_{i\bullet}}(X_{ij} - \mu(\Theta - i))\right)^2 \mid \Theta_i\right]\right\} \\ &= E\left[\frac{\sigma^2(\Theta_i)}{w_{i\bullet}}\right] = \frac{\sigma^2}{w_{i\bullet}}. \end{aligned}$$
- procjenitelj povjerenja je težinski prosjek ta dva procjenitelja gdje su težine dodijeljene svakom sumandu proporcionalne inverzu njegovog kvadratnog gubitka, tj.
$$\widehat{\widehat{\mu(\Theta_i)}} = \alpha_i \overline{X}_{i\bullet} + (1 - \alpha_i)\mu_0$$
gdje je
$$\alpha_i = \frac{\left(\frac{\sigma^2}{w_{i\bullet}}\right)^{-1}}{(\tau^2)^{-1} + \left(\frac{\sigma^2}{w_{i\bullet}}\right)^{-1}}.$$

Teorem 4.3.1. *Uz pretpostavke Bülhmann-Straubovog modela kvadratni gubitak procjenitelja povjerenja 4.3 dan je relacijom*
$$E\left[\left(\widehat{\widehat{\mu(\Theta_i)}} - \mu(\Theta_i)\right)^2\right] = (1 - \alpha_i)\,\tau^2 = \alpha_i\,\frac{\sigma^2}{w_{i\bullet}}. \tag{4.4}$$

Dokaz. Dokaz je analogan dokazu teorema 3.1.2. □

4.4 Homogeni procjenitelj povjerenja

Teorem 4.4.1 (homogeni procjenitelj povjerenja)**.** *Homogeni procjenitelj povjerenja od $\mu(\Theta_i)$ u Bülhmann-Straubovom modelu (uz pretpostavke Bülhmann-Straubovog modela) je dan s*

$$\widehat{\widehat{\mu(\Theta_i)}}^{hom} = \alpha_i \overline{X}_{i\bullet} + (1 - \alpha_i)\widehat{\widehat{\mu_0}} = \widehat{\widehat{\mu_0}} + \alpha_i(\overline{X}_{i\bullet} - \widehat{\widehat{\mu_0}}) \tag{4.5}$$

gdje je

$$\widehat{\widehat{\mu_0}} := \sum_{i=1}^{I} \frac{\alpha_i}{\alpha_\bullet} \overline{X}_{i\bullet}, \quad \alpha_i := \frac{w_{i\bullet}}{w_{i\bullet} + \frac{\sigma^2}{\tau^2}} \quad i \quad \alpha_\bullet := \sum_{i=1}^{I} \alpha_i.$$

Dokaz. Vidjeti u knjizi [1]. □

Napomena

- Uočimo da za procjenitelj μ_0 nismo uzeli težinski prosjek,

$$\overline{X} = \sum_{i=1}^{I} \frac{w_{i\bullet}}{w_{\bullet\bullet}} \overline{X}_{i\bullet}, \quad \text{gdje je} \quad w_{\bullet\bullet} = \sum_{ij} w_{ij}$$

  nego težinski prosjek povjerenja

$$\widehat{\widehat{\mu_0}} = \sum_{i=1}^{I} \frac{\alpha_i}{\alpha_\bullet} \overline{X}_{i\bullet}.$$

- U praksi, homogeni procjenitelj povjerenja je jednako važan, ako ne i važniji, od nehomogenog procjenitelja povjerenja. To je zbog toga što je procjena od μ_0 automatski ugrađena u formulu homogenog procjenitelja povjerenja, dok kod nehomogenog procjenitelja povjerenja to nije slučaj.

Kvadratni gubitak homogenog procjenitelja povjerenja dan je sljedećim teoremom:

Teorem 4.4.2. *Uz pretpostavke Bülhmann-Straubovog modela, kvadratni gubitak homogenog procjenitelja povjerenja 4.5 dan je s*

$$E\left[\left(\widehat{\widehat{\mu(\Theta_i)}}^{hom} - \mu(\Theta_i)\right)^2\right] = \tau^2(1 - \alpha_i)\left(1 + \frac{1 - \alpha_i}{\alpha_\bullet}\right) \tag{4.6}$$

Dokaz. Vidjeti u knjizi [1]. □

4.5 Procjena strukturnih parametara α^2 i τ^2 procjenitelja povjerenja

Vidjeli smo da formula za (nehomogeni) procjenitelj povjerenja 4.4 sadrži tri strukturna parametra μ_0, σ^2 i τ^2. Procjena od μ_0 je ugrađena u formulu homogenog procjenitelja povjerenja 4.5, stoga nam još ostaje odrediti parametre σ^2 i τ^2. U praksi, ta dva parametra su nepoznata i određuju se iz podataka za kolektiv.

Procjena parametra σ^2:

Za procjenu $\sigma^2(\Theta_i)$ iz uzorka promatramo sljedeći procjenitelj

$$S_i := \frac{1}{n-1} \sum_{j=1}^{n} w_{ij} (X_{ij} - \overline{X}_{i\bullet})^2.$$

Ako S_i zapišemo kao

$$\begin{aligned} S_i &= \frac{1}{n-1} \sum_{j=1}^{n} w_{ij} (X_{ij} - \mu(\Theta_i) + \mu(\Theta_i) - \overline{X}_{i\bullet})^2 \\ &= \frac{1}{n-1} \left\{ \sum_{j=1}^{n} w_{ij} (X_{ij} - \mu(\Theta_i))^2 - w_{i\bullet} (\mu(\Theta_i) - \overline{X}_{i\bullet})^2 \right\}, \end{aligned}$$

odmah vidimo da je

$$E[S_i \mid \Theta_i] = \sigma^2(\Theta_i),$$

stoga imamo

$$E[S_i] = E\{E[S_i \mid \Theta_i]\} = E[\sigma^2(\Theta_i)] = \sigma^2.$$

Dakle procjenitelj za σ^2 je

$$\widehat{\sigma^2} = \frac{1}{I} \sum_{i=1}^{I} \frac{1}{n-1} \sum_{j=1}^{n} w_{ij} (X_{ij} - \overline{X}_{i\bullet})^2 \tag{4.7}$$

i on je nepristran procjenitelj.

Procjena parametra τ^2:

a) Pretpostavimo da su ukupne težine za sve rizike jednake, tj. $w_{1\bullet} = w_{2\bullet} = ... = w_{I\bullet}$.

   Promatramo

$$T = \frac{1}{I-1} \sum_{i=1}^{I} (\overline{X}_{i\bullet} - \overline{X}_{\bullet})^2,$$

gdje je

$$\overline{X_\bullet} = \frac{1}{I} \sum_{i=1}^{I} \overline{X}_{i\bullet}.$$

Budući da je

$$E[T] = Var[\overline{X}_{i\bullet}] = \frac{\sigma^2}{w_{i\bullet}} + \tau^2,$$

dobijemo da je korigirani procjenitelj

$$\widehat{\widehat{\tau}}^2 = \frac{1}{I-1} \sum_{i=1}^{I} (\overline{X}_{i\bullet} - \overline{X_\bullet})^2 - \frac{\widehat{\sigma}^2}{w_{i\bullet}} \tag{4.8}$$

nepristarni procjenitelj za τ^2.

b) Općeniti slučaj:

Definiramo

$$T = \frac{I}{I-1} \sum_{i=1}^{I} \frac{w_{i\bullet}}{w_{\bullet\bullet}} (\overline{X}_{i\bullet} - \overline{X})^2,$$

gdje je

$$\overline{X} = \sum_{i=1}^{I} \frac{w_{i\bullet}}{w_{\bullet\bullet}} \overline{X}_{i\bullet}.$$

Nakon malo računanja dobijemo sljedeći nepristrani procjenitelj za τ^2:

$$\widehat{\widehat{\tau}}^2 = c \left\{ \frac{I}{I-1} \sum_{i=1}^{I} \frac{w_{i\bullet}}{w_{\bullet\bullet}} (\overline{X}_{i\bullet} - \overline{X})^2 - \frac{I\,\widehat{\sigma}^2}{w_{\bullet\bullet}} \right\}, \tag{4.9}$$

gdje je

$$c := \left\{ \frac{I-1}{I} \sum_{i=1}^{I} \frac{w_{i\bullet}}{w_{\bullet\bullet}} \left(1 - \frac{w_{i\bullet}}{w_{\bullet\bullet}}\right) \right\}^{-1}.$$

Napomena

- Uočimo da vrijedi

$$c = \begin{cases} 1, & \text{ako je } w_{1\bullet} = w_{2\bullet} = ... = w_{I\bullet} \\ > 1, & \text{inače.} \end{cases}$$

- Iako je procjenitelj $\widehat{\widehat{\tau}}^2$ nepristran, u praksi nije korišten jer ponekad može poprimiti negativne vrijednosti. U tom slučaju stavljamo $\widehat{\widehat{\tau}}^2 = 0$. Stoga τ^2 zapravo procjenjujemo koristeći sljedeći pristrani procjenitelj

$$\widehat{\tau}^2 = \max\left(\widehat{\widehat{\tau}}^2,\ 0\right), \tag{4.10}$$

gdje je $\widehat{\widehat{\tau}}^2$ dan sa formulom 4.9.

4.6 Empirijski procjenitelj povjerenja

Sada kada smo našli procjenitelje za strukturne parametre tražimo **empirijski procjenitelj povjerenja**. Dobijemo ga iz formule homogenog procjenitelja povjerenja 4.5 zamjenom strukturnih parametara σ^2 i τ^2 njihovim procjeniteljima.

Dakle, empirijski procjenitelj povjerenja dan je sa

$$\widehat{\widehat{\mu(\Theta_i)}}^{hom} = \widehat{\alpha}_i \overline{X}_{i\bullet} + (1 - \widehat{\alpha}_i)\widehat{\widehat{\mu}}_0$$

gdje su

$$\widehat{\alpha}_i = \frac{w_{i\bullet}}{w_{i\bullet} + \widehat{\kappa}}.\quad \widehat{\kappa} = \frac{\widehat{\sigma}^2}{\widehat{\tau}^2}.\quad \widehat{\widehat{\mu}}_0 = \frac{\sum_i \widehat{\alpha}_i \overline{X}_{i\bullet}}{\sum_i \widehat{\alpha}_i}.$$

Napomena

Uočimo da je $\widehat{\alpha}_i = 0$, ako je $\widehat{\tau}^2$. U tom slučaju $\widehat{\widehat{\mu}}_0$ je definiran sa

$$\widehat{\widehat{\mu}}_0 = \sum_i \frac{w_{i\bullet}}{w_{\bullet\bullet}} \overline{X}_{i\bullet}.$$

Primjer 4.6.1. *Sljedeća tablica sadrži podatke za portfelj osiguranja od požara koji se sastoji od pet grupa rizika, gdje su*

w – *osigurana vrijednost u milijunina,*

S – *ukupni iznos štete u tisućama,*

$X = \dfrac{S}{w}$ – *ukupni iznos štete po jedinici osigurane vrijednosti.*

grupe rizika \ godine		1	2	3	4	5	ukupno
grupa 1	w	729	786	872	951	1 019	4 357
	S	583	1 100	262	837	1 630	4 412
	X	0.80	1.40	0.30	0.88	1.60	1.01
grupa 2	w	1631	1802	2090	2300	2368	10191
	S	99	1 298	326	463	895	3081
	X	0.06	0.72	0.16	0.20	0.38	0.30
grupa 3	w	796	827	874	917	944	4358
	S	1433	496	699	1742	1038	5408
	X	1.80	0.60	0.80	1.90	1.10	1.24
grupa 4	w	3152	3454	3715	3859	4198	18378
	S	1765	4145	3121	4129	3358	16518
	X	0.56	1.20	0.84	1.07	0.80	0.90
grupa 5	w	400	420	422	424	440	2106
	S	40	0	169	1018	44	1271
	X	0.10	0.00	0.40	2.40	0.10	0.60
ukupno	w	6708	7289	7973	8451	8969	39390
	S	3920	7039	4577	8189	6965	30690
	X	0.58	0.97	0.57	0.97	0.78	0.78

Za svaku grupu rizika $i = 1, 2, ..., 5$ želimo procijeniti individualnu premiju $\mu(\Theta_i)$.

Trebamo izračunati:

a) *homogeni procjenitelj povjerenja koji se temelji na pretpostavkama Bülhmann -Straubovog modela (pretpostavke BS1 i BS2) uz zadani koeficijent povjerenja $\kappa = 300$,*

b) *isto kao pod a), ali uz zadani koeficijent povjerenja $\kappa = 600$,*

b) *isto kao pod a), ali koeficijent povjerenja $\kappa = \frac{\sigma^2}{\tau^2}$ trebamo procijeniti iz podataka.*

***Rješenje**:*

Homogeni procjenitelj povjerenja $\widehat{\widehat{\mu(\Theta)}}^{hom}$ predstavlja procjenu ukupnog iznosa štete po jedinici osigurane vrijednosti za šestu godinu, a $w \cdot \widehat{\widehat{\mu(\Theta)}}^{hom}$ predstvlja očekivani iznos štete za šestu godinu.

a) i b)
Da bismo izračunali homogeni procjenitelj povjerenja

$$\widehat{\widehat{\mu(\Theta_i)}}^{hom} = \alpha_i \overline{X}_{i\bullet} + (1 - \alpha_i)\widehat{\widehat{\mu_0}}$$

prvo izračunamo težine

$$\alpha_i = \frac{w_{i\bullet}}{w_{i\bullet} + \kappa}, \quad i = 1, 2, \ldots, 5,$$

pomoću kojih za svaku grupu rizika procijenimo individualnu premiju $\mu(\Theta_i), i = 1, 2, \ldots, 5$ *sa kolektivnom premijom*

$$\widehat{\widehat{\mu_0}} = \sum_{i=1}^{5} \frac{\alpha_i}{\alpha_\bullet} \overline{X}_{i\bullet}, \quad \textit{gdje je } \alpha_\bullet = \sum_{i=1}^{5} \alpha_i.$$

Ostale veličine koje su nam potrebne zadane su u tablici.

Za dani koeficijent povjerenja $\kappa = 3000$ *računamo,*

$$\begin{aligned}
\alpha_1 &= \frac{4357}{4357 + 3000} = 0.59,\\
\alpha_2 &= \frac{10191}{10191 + 3000} = 0.77,\\
\alpha_3 &= \frac{4358}{4358 + 3000} = 0.59,\\
\alpha_4 &= \frac{18378}{18378 + 3000} = 0.86,\\
\alpha_5 &= \frac{2106}{2106 + 3000} = 0.41.
\end{aligned}$$

Odavde dobijemo

$$\widehat{\widehat{\mu_0}} = \frac{0.59 * 1.01 + 0.77 * 0.30 + 0.59 * 1.24 + 0.86 * 0.90 + 0.41 * 0.60}{0.59 + 0.77 + 0.59 + 0.86 + 0.41} = \frac{2.58}{3.22} = 0.80.$$

Sada možemo izračunati homogene procjenitelje povjerenja za svaku godinu i sljedeću šestu godinu.

$$\begin{aligned}
\widehat{\widehat{\mu(\Theta_1)}}^{hom} &= 0.59 * 1.01 + (1 - 0.59) * 0.80 = 0.92,\\
\widehat{\widehat{\mu(\Theta_2)}}^{hom} &= 0.77 * 0.30 + (1 - 0.30) * 0.80 = 0.42,\\
\widehat{\widehat{\mu(\Theta_3)}}^{hom} &= 0.59 * 1.24 + (1 - 0.59) * 0.80 = 1.06,\\
\widehat{\widehat{\mu(\Theta_4)}}^{hom} &= 0.86 * 0.90 + (1 - 0.86) * 0.80 = 0.89,\\
\widehat{\widehat{\mu(\Theta_5)}}^{hom} &= 0.41 * 0.60 + (1 - 0.41) * 0.80 = 0.72.
\end{aligned}$$

Još nam je ostalo izračunati očekivane iznose štete

$$
\begin{aligned}
w_1 \cdot \widehat{\widehat{\mu(\Theta_1)}}^{hom} &= 4357 * 0.92 = 4008,\\
w_2 \cdot \widehat{\widehat{\mu(\Theta_2)}}^{hom} &= 10191 * 0.42 = 4280,\\
w_3 \cdot \widehat{\widehat{\mu(\Theta_3)}}^{hom} &= 4358 * 1.86 = 4620,\\
w_4 \cdot \widehat{\widehat{\mu(\Theta_4)}}^{hom} &= 18378 * 0.89 = 16356,\\
w_5 \cdot \widehat{\widehat{\mu(\Theta_5)}}^{hom} &= 2106 * 0.72 = 1516.
\end{aligned}
$$

Dakle, za $\kappa = 3000$ *imamo sljedeće rezultate:*

grupa rizika	1	2	3	4	5	ukupno
α_i	0.59	0.77	0.59	0.86	0.41	
$\widehat{\widehat{\mu}}_0$	0.80	0.80	0.80	0.80	0.80	
$\widehat{\widehat{\mu(\Theta_i)}}^{hom}$	0.92	0.42	1.06	0.89	0.72	0.78
$w_i \cdot \widehat{\widehat{\mu(\Theta_i)}}^{hom}$	4008	4280	4620	16356	1516	30780

Na temelju podataka za prethodnih pet godina izračunali smo da je procjena ukupnog iznosa štete po jedinici osigurane vrijednosti za šestu godinu 0.78, a očekivani iznos štete za šestu godinu iznosi 30780.

Da bismo vidjeli kako koeficijent povjerenja κ *utječe na promjenu rezultata, nad istim podacima proveli smo račun za* $\kappa = 6000$ *i dobili smo sljedeće rezultate:*

grupa rizika	1	2	3	4	5	ukupno
α_i	0.42	0.63	0.42	0.75	0.26	
$\widehat{\widehat{\mu}}_0$	0.80	0.80	0.80	0.80	0.80	
$\widehat{\widehat{\mu(\Theta_i)}}^{hom}$	0.89	0.48	0.98	0.87	0.74	0.78
$w_i \cdot \widehat{\widehat{\mu(\Theta_i)}}^{hom}$	3878	4892	4271	15989	1558	30588

Na temelju podataka za prethodnih pet godina izračunali smo da je procjena ukupnog iznosa štete po jedinici osigurane vrijednosti za šestu godinu 0.78, a očekivani iznos štete za šestu godinu iznosi 30588.

Dakle, povećanjem koeficijenta povjerenja dolazi do zanemarivog smanjenja očekivanog iznosa štete za sljedeću godinu.

c)

U ovom dijelu zadatka prvo trebamo procijeniti koeficijent povjerenja $\widehat{\kappa} = \frac{\widehat{\sigma}^2}{\widehat{\tau}^2}$*, odnosno trebamo izračunati* $\widehat{\sigma}^2$ *i* $\widehat{\tau}^2$*.*

Prvo računamo $\widehat{\sigma}^2$*:*

$$\begin{aligned}\widehat{\sigma^2} &= \frac{1}{5}\sum_{i=1}^{5}\frac{1}{5-1}\sum_{j=1}^{5} w_{ij}\,(X_{ij} - \overline{X}_{i\bullet})^2\\ &= \frac{1}{20}\left\{\begin{matrix} 729(0.80-1.01)^2 + 786(1.40-1.01)^2 + 872(0.30-1.01)^2\\ +951(0.88-1.01)^2 + 1019(1.60-1.01)^2 + 1631(0.06-0.30)^2\\ +1802(0.72-0.30)^2 + 2090(0.16-0.30)^2 + 2300(0.20-0.30)^2\\ +2368(0.38-0.30)^2 + 796(1.80-1.24)^2 + 827(0.60-1.24)^2\\ +874(0.80-1.24)^2 + 917(1.90-1.24)^2 + 944(1.10-1.24)^2\\ +3152(0.56-0.90)^2 + 3454(1.20-0.90)^2 + 3715(0.84-0.90)^2\\ +3859(1.07-0.90)^2 + 4198(0.80-0.90)^2 + 400(0.10-0.60)^2\\ +420(0.00-0.60)^2 + 422(0.40-0.60)^2 + 424(2.40-0.60)^2\\ +440(0.10-0.60)^2\end{matrix}\right\}\\ &= \frac{962+491+1176+842+1752}{20} = 261\end{aligned}$$

Računamo $\widehat{\tau}^2$*:*

Budući da težine po grupama nisu jednake koristim formule za računanje procjenitelja od τ^2 *u općem slučaju.*

Dakle,

$$\widehat{\tau}^2 = \max\left(\widehat{\widehat{\tau}}^2,\ 0\right)$$

gdje je

$$\widehat{\widehat{\tau}}^2 = c\left\{\frac{5}{5-1}\sum_{i=1}^{5}\frac{w_{i\bullet}}{w_{\bullet\bullet}}(\overline{X}_{i\bullet} - \overline{X})^2 - \frac{I\,\widehat{\sigma}^2}{w_{\bullet\bullet}}\right\},$$

a konstante c i $\overline{X}$ *računamo formulama:*

$$c := \left\{\frac{5-1}{5}\sum_{i=1}^{5}\frac{w_{i\bullet}}{w_{\bullet\bullet}}\left(1 - \frac{w_{i\bullet}}{w_{\bullet\bullet}}\right)\right\}^{-1}$$

i

$$\overline{X} = \sum_{i=1}^{5} \frac{w_{i\bullet}}{w_{\bullet\bullet}} \overline{X}_{i\bullet}.$$

Računamo:

$$\overline{X} = \frac{4357 * 1.01 + 10191 * 0.30 + 4358 * 1.24 + 18378 * 0.90 + 2106 * 0.60}{39390} = 0.78,$$

označimo sa T sljedeću jednadžbu:

$$\begin{aligned} T &= \frac{5}{5-1} \sum_{i=1}^{5} \frac{w_{i\bullet}}{w_{\bullet\bullet}} (\overline{X}_{i\bullet} - \overline{X})^2 \\ &= \frac{5}{4 * 39390} \begin{Bmatrix} 4357 * (1.01 - 0.78)^2 + 10191 * (0.30 - 0.78)^2 + 4358 * (1.2 - 0.78)^2 \\ + 18378 * (0.90 - 0.78)^2 + 2106 * (0.60 - 0.78)^2 \end{Bmatrix} \\ &= 0.12 \end{aligned}$$

i za c dobijemo sljedeći iznos

$$c = \left[\frac{4*}{5 * 39390^2} \begin{pmatrix} 4357 * 35033 + 10191 * 29199 + 4358 * 35032 \\ + 18378 * 21012 + 2106 * 37284 \end{pmatrix} \right]^{-1} = 1.81.$$

Sada imamo da je

$$\widehat{\widehat{\tau}}^2 = c * \{T - \frac{5 * \widehat{\sigma}^2}{39390}\} = 1.81 * (0.12 - 0.03) = 0.16.$$

Odavde slijedi

$$\widehat{\tau}^2 = \max\left(\widehat{\widehat{\tau}}^2,\ 0\right) = 0.16.$$

Dakle, koeficijent provjerenja je $\kappa = \frac{261}{0.16} = 1631$.
I sada opet sve računamo na isti način kao u a) i b) dijelu zadataka.
Dobili smo sljedeće rezultate:

grupa rizika	1	2	3	4	5	ukupno
α_i	0.73	0.86	0.73	0.92	0.56	
$\widehat{\widehat{\mu}_0}$	0.80	0.80	0.80	0.80	0.80	
$\widehat{\widehat{\mu(\Theta_i)}}^{hom}$	0.95	0.37	1.12	0.89	0.69	0.78
$w_i \cdot \widehat{\widehat{\mu(\Theta_i)}}^{hom}$	4139	3771	4881	16356	1453	30600

Bibliografija

[1] Hans Bülhmann, Alois Gisler, *A Course in Credibility Theory and its Applications*, Springer, Berlin, 2005.

[2] Nikola Sarapa, *Teorija vjerojatnosti*, Školska knjiga, Zagreb, 2002.

[3] *SUBJECT C1 CORE READING*, HAD, Zagreb, 1996.

Sažetak

U ovom diplomskom radu bavili smo se određivanjem što točnije individualne premije za rizik. Budući da je konkurencija na tržištu velika, cilj svakog osiguravajućeg društva je ponuditi što točniju premiju za svaki rizik. Premije ne smiju biti previsoke jer bi odbile klijente, a ne smiju biti ni preniske kako bi osiguravajuće društvo skupilo dovoljno sredstava kojima bi u svakom trenutku moglo ispuniti obveze prema svojim klijentima.

Individualna premija temelji se na podacima o štetama pojedinaca u prethodnim periodima. Tu dolazimo do problema procjene premije za osiguranika koji je u tom osiguranju kratko vrijeme ili je novi osiguranik. U tom slučaju premiju bolje možemo procijeniti s prosječnom štetom sličnih rizika u osiguranju. Nju smo nazvali kolektivna premija. Ona je dobra za posve novog osiguranika, ali jasno je da za ostale osiguranike premija treba ovisiti i o individualnoj i o kolektivnoj premiji.

Pokazali smo da je najbolja procjena premije Bayesova premija koja ovisi i o ponašanju pojedinca i o prosječnom ponašanju sličnih rizika. Međutim, Bayesov procjenitelj je često puta komplicirano egzaktno izračunati. Štoviše, da bismo izračunali Bayesov procjenitelj moramo odrediti uvjetne distribucije kao i *a priori* distribucije šteta, što u praksi često ne možemo zaključiti niti iz danih podataka niti pogoditi intuicijom. Stoga smo umjesto njega procjenitelj za individualnu premiju tražili među linearnim funkcijama našeg uzorka i takve procjenitelje smo zvali procjenitelji povjerenja.

Procjenitelji povjerenja imaju oblik konveksne kombinacije prosječne vrijednosti podataka za individualni rizik i prosječne vrijednosti podataka za cijeli kolektiv. U kojoj mjeri na procjenu premije utječe koja od tih dviju varijabli, to je predmet kojim se bavi teorija povjerenja. Što je uzorak podataka veći to će premija biti bliža prosječnoj vrijednosti podataka za individualni rizik, a za manji uzorak premija će biti bliža prosječnoj vrijednosti podataka za cijeli kolektiv.

Summary

In this thesis we deal with determining the precise individual risk premium. Since the competition in the market is high, the goal of every insurance company is to offer the most accurate premium for each risk. Premiums should not be too high because that can refuse customers, on the other hand premiums should not be too low because the insurance company needs to gather enough funds to fulfill obligations to their clients.

Individual premium is based on individual claims data in prior periods. Here we come to the problem of how to estimate premium for the insured who is in the insurance for a short period of time or a new insured. In this case, the premium can better estimate the average damage similar to insurance risk. We called it the collective premium. It is good for the insured who is entirely new, but it is clear that for other insured premium it should depend on the individual and the collective premium.

We have shown that the best estimate of the premium is Bayesian premium that depends on the behavior of individuals and the average behavior of similar risk. However, the Bayesian estimator is often complicated to calculate exactly. Moreover, in order to calculate the Bayesian estimator we have to specify the conditional distribution as well as the a priori distribution for damage, which, in practice, can often neither be inferred from any given data nor guessed by intuition. So instead the Bayesian estimator we look for the best estimator in the class of all linear estimator functions of our sample and such estimators are called credibility estimators.

Credibility estimators have convex form as combinations of the average value of the individual risk data and average value of the entire collective data. The question is, which of these two variables affects the premium more, this question is a subject in credibility theory. If the sample of data is larger, premiums will be closer to the average value of the individual risk data, and for smaller sample premium will be closer to the average value of the whole collective data.

Životopis

Rođena sam 25. travnja 1988. godine u Virovitici. Nakon završene osnovne škole Ivane Brlić-Mažuranić, 2003. godine u Virovitici upisujem gimnaziju Petra Preradovića, matematički smjer. Godine 2007. upisujem Preddiplomski sveučilišni studij Matematika na Matematičkom odsjeku Prirodoslovno-matematičkog fakulteta u Zagrebu. Nakon završenog preddiplomskog studija, 2012. godine, upisujem Diplomski sveučilišni studij Matematička statistika na istom fakultetu.

Godine 2011. kada se na Prirodoslovno-matematičkom fakultetu osnovao ogranak studentske udruge eSTUDENT, postajem članica Tima za prakse i pripravništva. Od 2011. do 2014. godine pohađam SOVA jezično učilište učeći engleski jezik. Akademske godine 2013/14. odslušala sam tečajeve iz primijenjene statistike i programskog sustava SAS-a u sveučilišnom računarskom centru SRCE.